PCT

(22) International Filing Date:

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ :		(11) International Publication Number:	WO 00/58473
C12N 15/12, C07K 14/47, 16/18, G01N 33/566, C12Q 1/68, C12N 15/11, 15/62, A01K 67/027, A61K 38/00	A2	(43	i) International Publication Date:	5 October 2000 (05.10.00)
(21) International Application Number: PCT/US	500/086	1211	(72) Inventors; and	A CHIMITETE Dishord A

31 March 2000 (31.03.00)

(30) Priority Data:	
60/127,607	31 March 1999 (31.03.99)
60/127 636	2 April 1999 (02.04.99)

 60/127,607
 31 March 1999 (31.03.99)
 US

 60/127,636
 2 April 1999 (02.04.99)
 US

 60/127,728
 5 April 1999 (05.04.99)
 US

 09/540,763
 30 March 2000 (30.03.00)
 US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/127,607 (CIP)
Filed on	31 March 1999 (31.03.99)
US	60/127,636 (CIP)
Filed on	2 April 1999 (02.04.99)
US	60/127,728 (CIP)
Filed on	5 April 1999 (05.04.99)
US	09/540,763 (CIP)
Filed on	30 March 2000 (30.03.00)

(71) Applicant (for all designated States except US): CURAGEN CORPORATION [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US).

- (75) Inventors/Applicants (for US only): SHIMKETS, Richard, A. [US/US]; 191 Leete Street, West Haven, CT 06516 (US). LEACH, Martin [GB/US]; 884 School Street, Webster, MA 01570 (US).
- (74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).
- (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: NUCLEIC ACIDS INCLUDING OPEN READING FRAMES ENCODING POLYPEPTIDES; "ORFX"

(57) Abstract

The present invention provides open reading frames ORFX, encoding isolated polypeptides, as well as polynucleotides encoding ORFX and antibodies that immunospecifically bind to ORFX or any derivative, variant, mutant, or fragment of the ORFX polypeptides, polynucleotides or antibodies. The invention additionally provides methods in which the ORFX polypeptide, polynucleotide and antibody are used in detection and treatment of a broad range of pathological states, as well as to other uses.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	$\mathbf{F}\mathbf{R}$	France	LU	Luxembourg	SN	Scnegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	ТJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
\mathbf{BF}	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	$\mathbf{U}\mathbf{Z}$	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	П	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY

5

BACKGROUND OF THE INVENTION

The invention relates generally to nucleic acids and polypeptides encoded thereby, and methods of using these nucleic acids and polypeptides.

10

SUMMARY OF THE INVENTION

The invention is based in part on the discovery of nucleic acids that include open reading frames encoding novel polypeptides, and on the polypeptides encoded thereby. The nucleic acids and polypeptides are collectively referred to herein as "ORFX".

Accordingly, in one aspect, the invention provides an isolated nucleic acid molecule (SEQ ID NO:2n-1, wherein n is an integer between 1-3161), that encodes novel polypeptide, or a fragment, homolog, analog or derivative thereof. The nucleic acid can include, e.g., a nucleic acid sequence encoding a polypeptide at least 85% identical to a polypeptide comprising the amino acid sequences of SEQ ID NO:2n, wherein n is an integer between 1-3161. The nucleic acid can be, e.g., a genomic DNA fragment, or a cDNA molecule.

20

25

15

Also included in the invention is a vector containing one or more of the nucleic acids described herein, and a cell containing the vectors or nucleic acids described herein.

The invention is also directed to host cells transformed with a recombinant expression vector comprising any of the nucleic acid molecules described above.

In another aspect, the invention includes a pharmaceutical composition that includes an ORFX nucleic acid and a pharmaceutically acceptable carrier or diluent.

In a further aspect, the invention includes a substantially purified ORF polypeptide, *e.g.*, any of the ORFX polypeptides encoded by an ORFX nucleic acid, and fragments, homologs, analogs, and derivatives thereof. The invention also includes a pharmaceutical composition that includes a ORFX polypeptide and a pharmaceutically acceptable carrier or diluent.

5

10

15

In a still a further aspect, the invention provides an antibody that binds specifically to an ORFX polypeptide. The antibody can be, *e.g.*, a monoclonal or polyclonal antibody, and fragments, homologs, analogs, and derivatives thereof. The invention also includes a pharmaceutical composition including ORFX antibody and a pharmaceutically acceptable carrier or diluent. The invention is also directed to isolated antibodies that bind to an epitope on a polypeptide encoded by any of the nucleic acid molecules described above.

The invention also includes kits comprising any of the pharmaceutical compositions described above.

The invention further provides a method for producing an ORFX polypeptide by providing a cell containing a ORFX nucleic acid, e.g., a vector that includes a ORFX nucleic acid, and culturing the cell under conditions sufficient to express the ORFX polypeptide encoded by the nucleic acid. The expressed ORFX polypeptide is then recovered from the cell. Preferably, the cell produces little or no endogenous ORFX polypeptide. The cell can be, e.g., a prokaryotic cell or eukaryotic cell.

20

The invention is also directed to methods of identifying an ORFX polypeptide or nucleic acids in a sample by contacting the sample with a compound that specifically binds to the polypeptide or nucleic acid, and detecting complex formation, if present.

The invention further provides methods of identifying a compound that modulates the activity of a ORFX polypeptide by contacting ORFX polypeptide with a compound and determining whether the ORFX polypeptide activity is modified.

25

The invention is also directed to compounds that modulate ORFX polypeptide activity identified by contacting a ORFX polypeptide with the compound and determining whether the compound modifies activity of the ORFX polypeptide, binds to the ORFX polypeptide, or binds to a nucleic acid molecule encoding a ORFX polypeptide.

30

In a another aspect, the invention provides a method of determining the presence of or predisposition of an ORFX-associated disorder in a subject. The method includes providing a sample from the subject and measuring the amount of ORFX polypeptide in the subject sample.

The amount of ORFX polypeptide in the subject sample is then compared to the amount of ORFX polypeptide in a control sample. An alteration in the amount of ORFX polypeptide in the subject protein sample relative to the amount of ORFX polypeptide in the control protein sample indicates the subject has a tissue proliferation-associated condition. A control sample is preferably taken from a matched individual, *i.e.*, an individual of similar age, sex, or other general condition but who is not suspected of having a tissue proliferation-associated condition. Alternatively, the control sample may be taken from the subject at a time when the subject is not suspected of having a tissue proliferation-associated disorder. In some embodiments, the ORFX is detected using a ORFX antibody.

5

10

15

20

25

30

In a further aspect, the invention provides a method of determining the presence of or predisposition of an ORFX-associated disorder in a subject. The method includes providing a nucleic acid sample, e.g., RNA or DNA, or both, from the subject and measuring the amount of the ORFX nucleic acid in the subject nucleic acid sample. The amount of ORFX nucleic acid sample in the subject nucleic acid is then compared to the amount of an ORFX nucleic acid in a control sample. An alteration in the amount of ORFX nucleic acid in the sample relative to the amount of ORFX in the control sample indicates the subject has a tissue proliferation-associated disorder.

In a still further aspect, the invention provides method of treating or preventing or delaying a ORFX-associated disorder. The method includes administering to a subject in which such treatment or prevention or delay is desired a ORFX nucleic acid, a ORFX polypeptide, or an ORFX antibody in an amount sufficient to treat, prevent, or delay a tissue proliferation-associated disorder in the subject.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

5

10

15

20

25

30

DETAILED DESCRIPTION OF THE INVENTION

The invention provides novel polypeptides and nucleotides encoded thereby. The polynucleotides and their encoded polypeptides can be grouped according to the functions played by their gene products. Such functions include, structural proteins, proteins from which associated with metabolic pathways fatty acid metabolism, glycolysis, intermediary metabolism, calcium metabolism, proteases, and amino acid metabolism, etc.

Included in the invention are 3161 novel nucleic acid sequences and their encoded polypeptides. The sequences are collectively referred to as "ORFX nucleic acids" or ORFX polynucleotides" and the corresponding encoded polypeptide is referred to as a "ORFX polypeptide" or ORFX protein". For example, an ORFX nucleic acid according to the invention is a nucleic acid including an ORF1 nucleic acid, and an ORF polypeptide according to the invention is a polypeptide that includes the amino acid sequence of an ORF1 polypeptide. Unless indicated otherwise, "ORFX" is meant to refer to any of the ORF1-3161 sequences disclosed herein.

Table 1 provides a summary of the ORFX nucleic acids and their encoded polypeptides are summarized in Table 1. Nucleic acid sequences and polypeptide sequences for ORFX nucleic acids according to the invention is provided in the section of the specification entitled "Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences."

Column 1 of Table 1, entitled "ORF #", denotes an ORF number assigned to a nucleic acid containing an open reading frame according to the invention.

Column 2 of Table 1, entitled "Internal Identification number (Nucleic Acid Sequence Identification Number, Polypeptide Sequence Identification Number), provides an internal identification number for the indicated ORF, along with sequence identification numbers (SEQ ID NOs.) corresponding to the indicated ORF. In general, for an ORFn according to the invention (wherein n is any integer from 1 to 3161), a nucleic acid corresponding to the ORF is SEQ ID NO:2n-1, and an amino acid sequence encoded by the ORF is SEQ ID NO:2n. For example, a nucleic acid sequence corresponding to an ORF1 nucleic acid is SEQ ID NO:1, and a polypeptide sequence corresponding to an ORF1 polypeptide is SEQ ID NO:2. Similarly, a

nucleic acid sequence corresponding to an ORF4 nucleic acid is SEQ ID NO:7, and a polypeptide sequence corresponding to an ORF4 polypeptide is SEQ ID NO:8; a nucleic acid sequence corresponding to an ORF198 nucleic acid sequence is SEQ ID NO:395, and a polypeptide sequence corresponding to an ORF198 polypeptide is SEQ ID NO:396. Nucleic acid sequences and polypeptide sequences for ORFX nucleic acids according to the invention are provided in the section of the specification entitled "Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences."

5

10

15

20

25

Column 2 of Table 1, entitled "Protein Similarity", lists previously described proteins that are related to polypeptides encoded by the ORFs. Genbank identifiers for the previously described proteins are provided. These can be retrieved from http://www.ncbi.nlm.nih.gov/.

To determine similarity to previously described proteins, polypeptides encoded by ORFX DNA sequences were tested using the Framesearch Algorithm against a nonredundant version of the GenPept Database from NCBI/Genbank. DNA sequences that had a score of '90' or above (Framesearch algorithm score, Edelman et. al. GCG Genetics) to a known protein were selected. Open reading frames were extended beyond the region of the protein matched using standard DNA translation and codon tables. Novel proteins that lacked a protein match were translated against the standard genetic codons and proteins with an ORF at least 80 amino acids and containing a Methionine start are included in the Table.

Column 3 of Table 3, entitled "Protein Domains", lists previously described protein domains, designated by pfam entries, that are present in polypeptides encoded by the ORFs, Also included in column 3 are proteins in which these domains are present. The pfam entries can be retrieved from http://pfam.wustl.edu/. DNA sequences were translated in all six frames and tested using the Hmmer Algorithm against the Pfam Database (References to the algorithm and Pfam database can be found at http://pfam.wustl.edu). Translated DNA sequences that matched a protein domain entry in the Pfam database AND had a score of 7.5' were selected.

Column 4 of Table 3, entitled "Protein Classification", lists the type of classification assigned for the protein, based on its homology. Examples of proteins in the classification include the following proteins:

Amylases

Amylase is responsible for endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides. Variations in amylase gene may be indicative of delayed maturation and of various amylase producing neoplasms and carcinomas.

5 Amyloid

10

15

20

25

The serum amyloid A (SAA) proteins comprise a family of vertebrate proteins that associate predominantly with high density lipoproteins (HDL). The synthesis of certain members of the family is greatly increased in inflammation. Prolonged elevation of plasma SAA levels, as in chronic inflammation, 15 results in a pathological condition, called amyloidosis, which affects the liver, kidney and spleen and which is characterized by the highly insoluble accumulation of SAA in these tissues. Amyloid selectively inhibits insulin-stimulated glucose utilization and glycogen deposition in muscle, while not affecting adipocyte glucose metabolism. Deposition of fibrillar amyloid proteins intraneuronally, as neurofibrillary tangles, extracellularly, as plaques and in blood vessels, is characteristic of both Alzheimer's disease and aged Down's syndrome. Amyloid deposition is also associated with type II diabetes mellitus.

Angiopoeitin

Members of the angiopoeitin/fibrinogen family have been shown to stimulate the generation of new blood vessels, inhibit the generation of new blood vessels, and perform several roles in blood clotting. This generation of new blood vessels, called angiogenesis, is also an essential step in tumor growth in order for the tumor to get the blood supply it needs to expand. Variation in these genes may be predictive of any form of heart disease, numerous blood clotting disorders, stroke, hypertension and predisposition to tumor formation and metastasis. In particular, these variants may be predictive of the response to various antihypertensive drugs and chemotherapeutic and anti-tumor agents.

Apoptosis-related proteins

Active cell suicide (apoptosis) is induced by events such as growth factor withdrawal and toxins. It is controlled by regulators, which have either an inhibitory effect on programmed cell

death (anti-apoptotic) or block the protective effect of inhibitors (pro-apoptotic). Many viruses have found a way of countering defensive apoptosis by encoding their own anti-apoptosis genes preventing their target-cells from dying too soon. Variants of apoptosis related genes may be useful in formulation of anti-aging drugs.

Cadherin, Cyclin, Polymerase, Oncogenes, Histones, Kinases

Members of the cell division/cell cycle pathways such as cyclins, many transcription factors and kinases, DNA polymerases, histones, helicases and other oncogenes play a critical role in carcinogenesis where the uncontrolled proliferation of cells leads to tumor formation and eventually metastasis. Variation in these genes may be predictive of predisposition to any form of cancer, from increased risk of tumor formation to increased rate of metastasis. In particular, these variants may be predictive of the response to various chemotherapeutic and anti-tumor agents.

Colony-stimulating factor-related proteins

Granulocyte/macrophage colony-stimulating factors are cytokines that act in hematopoiesis by controlling the production, differentiation, and function of 2 related white cell populations of the blood, the granulocytes and the monocytes-macrophages.

Complement-related proteins

5

10

15

20

25

Complement proteins are immune associated cytotoxic agents, acting in a chain reaction to exterminate target cells to that were opsonized (primed) with antibodies, by forming a membrane attack complex (MAC). The mechanism of killing is by opening pores in the target cell membrane. Variations in 20 complement genes or their inhibitors are associated with many autoimmune disorders. Modified serum levels of complement products cause edemas of various tissues, lupus (SLE), vasculitis, glomerulonephritis, renal failure, hemolytic anemia, thrombocytopenia, and arthritis. They interfere with mechanisms of ADCC (antibody dependent cell cytotoxicity), severely impair immune competence and reduce phagocytic ability. Variants of complement genes may also be indicative of type I diabetes mellitus, meningitis neurological disorders such as nemaline myopathy, neonatal hypotonia, muscular disorders such as congenital myopathy and other diseases.

Cytochrome

The respiratory chain is a key biochemical pathway which is essential to all aerobic cells. There are five different cytochromes involved in the chain. These are heme bound proteins which serve as electron carriers. Modifications in these genes may be predictive of ataxia areflexia, dementia and myopathic and neuropathic changes in muscles. Also, association with various types of solid tumors.

Kinesins

5

10

15

20

25

Kinesins are tubulin molecular motors that function to transport organelles within cells and to move chromosomes along microtubules during cell division. Modifications of these genes may be indicative of neurological disorders such as Pick disease of the brain, tuberous sclerosis.

Cytokines, Interferon, Interleukin

Members of the cytokine families are known for their potent ability to stimulate cell growth and division even at low concentrations. Cytokines such as erythropoietin are cell-specific in their growth stimulation; erythropoietin is useful for the stimulation of the proliferation of erythroblasts. Variants in cytokines may be predictive for a wide variety of diseases, including cancer predisposition.

G-protein coupled receptors

G-protein coupled receptors (also called R7G) are an extensive group of hormones, neurotransmitters, odorants and light receptors which transduce extracellular signals by interaction with guanine nucleotide-binding (G) proteins. Alterations in genes coding for G-coupled proteins may be involved in and indicative of a vast number of physiological conditions. These include blood pressure regulation, renal dysfunctions, male infertility, dopamine associated cognitive, emotional, and endocrine functions, hypercalcemia, chondrodysplasia and osteoporosis, pseudohypoparathyroidism, growth retardation and dwarfism.

Thioesterases

5

Eukaryotic thiol proteases are a family of proteolytic enzymes which contain an active site cysteine. Catalysis proceeds through a thioester intermediate and is facilitated by a nearby histidine side chain; an asparagine completes the essential catalytic triad. Variants of thioester associated genes may be predictive of neuronal disorders and mental illnesses such as Ceroid Lipoffiscinosis, Neuronal 1, Infantile, Santavuori disease and more.

The key to the molecule type is as follows:

10	Abbrev:	Title:
	amylase	amylase protein
	amylaseinhib	amylase inhibitor
	amyloid	amyloid protein
15	apoptosis	apoptosis associated protein
	apoptosisinhib	apoptosis inhibitors
	apoptosisrecep	apoptosis receptors
	ATPase_associated	ATPase associated protein
	biotindep	biotin dependent enzyme/protein
20	cadherin	cadherin protein
	calcium_channel	calcium channel protein
	carboxylase	carboxylase protein
	cathepsin	cathepsin/carboxypeptidases
	cathepsininhib	cathepsin/carboxypeptidase inhibitor
25	chloride_channel	chloride channel protein
	collagen	collagen
	complement	complement protein
	complementrecept	complement receptor protein
	complementinhib	complement inhibitor
30	csf	colony stimulating factor
	csfrecept	colony stimulating factor receptor
	cyclin	cyclin protein
	cyto450	cytochrome p450 protein
	cytochrome	cytochrome related protein
35	deaminase	deaminase
	dehydrogenase	dehydrogenase
	desaturase	desaturase
	dna_rna_bind	DNA/RNA binding protein/factor
	dna_rna_inhib	DNA/RNA binding protein/factor inhibitor
40	dynein	dynein

		•
	elastase	elastase
	elastaseinhib	elastase inhibitor
	eph	EPH family of tyrosine kinases
	esterase	esterase
5	esteraseinhib	esterase inhibitor
	fgf	fibroblast growth factor
	fgfreceptor	fibroblast growth factor receptor
	gaba	GABA receptor
	glucoamylase	glucoamylase
10	glucoronidase	glucoronidase
	glycoprotein	glycoprotein
	Guanylyl	guanylylate cyclase
	helicase	helicase
	histone	histone
15	HOM	homologous
	homeobox	homeobox protein
	hydrolase	hydrolase
	hydroxysteroid	hydroxysteroid associated protein
	hypoxanthine	hypoxanthine associated protein
20	immunoglob	immunoglobulin
	immunoglobrecept	immunoglobulin receptor
	interferon	interferon
	interleukin	interleukin
	interleukinrecept	interleukin receptor
25	isomerase	isomerase
	isomeraseinhibitor	isomerase inhibitor
	isomerasereceptor	isomerase receptor
	kinase	kinase
	kinaseinhibitor	kinase inhibitor
30	kinasereceptor	kinase receptor
	kinesin	kinesin
	laminin	laminin associated protein
	lipase	lipase
	metallothionein	metallothionein
35	MHC	major histocompatability complex
	misc_channel	miscellaneous channel
	ngf	nerve growth factor
	nuci_recpt	nuclear receptor
	nuclease	nuclease
40	oncogene	oncogene associated protein
	oxidase	oxidase
	oxygenase	oxygenase
	peptidase	peptidase
	peroxidase	peroxidase
45	phosphatase	phosphatase
	phosphataseinhib	phosphatase inhibitor

	phosphorylase PIR	phosphorylase PIR DATABASE (release 56, 29-OCT- 1998)
	polymerase	polymerase
5	potassium_channel	potassium channel protein
	prostaglandin	prostaglandin
	protease	protease
	proteaseinhib	protease inhibitor
	reductase	reductase
10	ribosomalprot	ribosomal associated protein
	RTR	EMBLDATABASE translated entries not to be incorporated into SWISS-PROT (20-JUL-1998)
	SIM	similar
15	SPTR	EMBL DATABASE translated entries to be incorporated into SWISS-PROT (20-JUL-1998)
	struct	structural associated protein
	sulfotransferase	sulfotransferase
20	SWP	SWISS-PROT DATABASE (release 18- OCT-1998)
	SWPN	SWISS-PROT Update (release 11-NOV-98)
	synthase	synthase
	tgf	transforming growth factor
25	tgfreceptor	transforming growth factor receptor
	thioesterase	thioesterase
	thiolase	thiolase
	tm7	seven transmembrane domain G-protein
		coupled receptor
30	tnf	necrosis factor receptor
	traffic	tumor necrosis factor
	tnfreceptor	tumor trafficking associated protein
	TRN	EMBL DATABASE translated entries update (20-JUL-1998)
35	transcriptfactor	transcription factor
	transferase	transferase
	transport	transport protein
	tubulin	tubulin
	ubiquitin	ubiquitin
40	unclassified	Protein not categorized into one of the
		aforementioned protein families
	water channel	water channel protein

Column 5 of Table 1, entitled, "Cells or Tissues in Which Gene is Expressed", denotes tissues, represented by five digit numbers, in which RNA homologous to the ORF nucleic acid sequences is present. Tissues or cells corresponding to the numbers are provided in Table 2.

ORFX nucleic acids, and their encoded polypeptides, according to the invention are useful in a variety of applications and contexts. For example, various ORFX nucleic acids and polypeptides according to the invention are useful, *inter alia*, as novel members of the protein families indicated in Table 1, and/or according to the presence of domains and sequence relatedness to previously described proteins as summarized in Table 1.

ORFX nucleic acids and polypeptides according to the invention can also be used to identify cell types listed in Table 1 for an indicated ORFX according to the invention.

Additional utilities for ORFX nucleic acids and polypeptides according to the invention are disclosed herein.

ORFX Nucleic Acids

5

10

15

20

25

The novel nucleic acids of the invention include those that encode an ORFX or ORFX-like protein, or biologically active portions thereof. The nucleic acids include nucleic acids encoding polypeptides that include the amino acid sequence of one or more of SEQ ID NO:2n, wherein n = 1 to 3161. The encoded polypeptides can thus include, e.g., the amino acid sequences of SEQ ID NO: 2, 4, 6, 8, 10, ..., 6310, 6312, 6314, 6316, 6318, 6320, and/or 6322.

In some embodiments, a nucleic acid encoding a polypeptide having the amino acid sequence of one or more of SEQ ID NO:2n (wherein n = 1 to 3161) includes the nucleic acid sequence of any of SEQ ID NO:2n-1 (wherein n = 1 to 3161), or a fragment thereof.

Additionally, the invention includes mutant or variant nucleic acids of any of SEQ ID NO:2n-1 (wherein n = 1 to 3161), or a fragment thereof, any of whose bases may be changed from the disclosed sequence while still encoding a protein that maintains its ORFX -like activities and physiological functions. The invention further includes the complement of the nucleic acid sequence of any of SEQ ID NO:2n-1 (wherein n = 1 to 3161), including fragments, derivatives,

analogs and homolog thereof. The invention additionally includes nucleic acids or nucleic acid fragments, or complements thereto, whose structures include chemical modifications.

Also included are nucleic acid fragments sufficient for use as hybridization probes to identify ORFX-encoding nucleic acids (e.g., ORFX mRNA) and fragments for use as polymerase chain reaction (PCR) primers for the amplification or mutation of ORFX nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA), RNA molecules (e.g., mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

5

10

15

20

25

30

"Probes" refer to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), 100 nt, or as many as about, e.g., 6,000 nt, depending on use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. Probes may be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

An "isolated" nucleic acid molecule is one that is separated from other nucleic acid molecules that are present in the natural source of the nucleic acid. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated ORFX nucleic acid molecule can contain less than about 50 kb, 25 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:2n-1 (wherein n=1 to 3161), or a complement of any of this nucleotide sequence, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of any of SEQ ID NO:2n-1 (wherein n=1 to 3161) as a hybridization probe, ORFX nucleic acid sequences can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook et~al., eds., Molecular Cloning: A Laboratory Manual 2^{nd} Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, et~al., eds., Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, 1993.)

5

10

15

20

25

30

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to ORFX nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having about 10 nt, 50 nt, or 100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at lease 6 contiguous nucleotides of any of SEQ ID NO:2n-1 (wherein n = 1 to 3161), or a complement thereof. Oligonucleotides may be chemically synthesized and may be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in any of SEQ ID NO:2n-1 (wherein n=1 to 3161). In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in any of SEQ ID NO:2n-1 (wherein n=1 to 3161), or a portion of this nucleotide sequence. A nucleic acid molecule that is complementary to the nucleotide sequence shown in

is one that is sufficiently complementary to the nucleotide sequence shown in of any of SEQ ID NO:2n-1 (wherein n=1 to 3161) that it can hydrogen bond with little or no mismatches to the nucleotide sequence shown in of any of SEQ ID NO:2n-1 (wherein n=1 to 3161),, thereby forming a stable duplex.

5

10

15

20

25

30

As used herein, the term "complementary" refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term "binding" means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, Von der Waals, hydrophobic interactions, etc. A physical interaction can be either direct or indirect. Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of any of SEQ ID NO:2n-1 (wherein n=1 to 3161), e.g., a fragment that can be used as a probe or primer, or a fragment encoding a biologically active portion of ORFX. Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively, and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type.

Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the

invention, in various embodiments, by at least about 70%, 80%, 85%, 90%, 95%, 98%, or even 99% identity (with a preferred identity of 80-99%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. See *e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below. An exemplary program is the Gap program (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison, WI) using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2: 482-489, which is incorporated herein by reference in its entirety).

5

10

15

20

25

30

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of ORFX polypeptide. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the present invention, homologous nucleotide sequences include nucleotide sequences encoding for a ORFX polypeptide of species other than humans, including, but not limited to, mammals, and thus can include, e.g., mouse, rat, rabbit, dog, cat cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the nucleotide sequence encoding human ORFX protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in any of SEQ ID NO:2n (wherein n = 1 to 3161) as well as a polypeptide having ORFX activity. Biological activities of the ORFX proteins are described below. A homologous amino acid sequence does not encode the amino acid sequence of a human ORFX polypeptide.

The nucleotide sequence determined from the cloning of the human ORFX gene allows for the generation of probes and primers designed for use in identifying the cell types disclosed and/or cloning ORFX homologues in other cell types, e.g., from other tissues, as well as ORFX homologues from other mammals. The probe/primer typically comprises a substantially purified

oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 or more consecutive sense strand nucleotide sequence of SEQ ID NO:2n-1 (wherein n=1 to 3161); or an anti-sense strand nucleotide sequence of SEQ ID NO:2n-1 (wherein n=1 to 3161); or of a naturally occurring mutant of SEQ ID NO:2n-1 (wherein n=1 to 3161).

Probes based on the human ORFX nucleotide sequence can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a ORFX protein, such as by measuring a level of a ORFX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting ORFX mRNA levels or determining whether a genomic ORFX gene has been mutated or deleted.

"A polypeptide having a biologically active portion of ORFX" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically active portion of ORFX" can be prepared by isolating a portion of SEQ ID NO:2n-1 (wherein n = 1 to 3161), that encodes a polypeptide having a ORFX biological activity (biological activities of the ORFX proteins are summarized in Table 1), expressing the encoded portion of ORFX protein (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of ORFX. For example, a nucleic acid fragment encoding a biologically active portion of ORFX can optionally include a domain as shown in Table 1, column 4.

ORFX variants

5

10

15

20

25

30

The invention further encompasses nucleic acid molecules that differ from the disclosed ORFX nucleotide sequences due to degeneracy of the genetic code. These nucleic acids thus encode the same ORFX protein as that encoded by the nucleotide sequence shown in SEQ ID NO:2n-1 (wherein n=1 to 3161). In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in any of SEQ ID NO:2n (wherein n=1 to 3161).

In addition to the human ORFX nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of ORFX may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the ORFX gene may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a ORFX protein, preferably a mammalian ORFX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the ORFX gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in ORFX that are the result of natural allelic variation and that do not alter the functional activity of ORFX are intended to be within the scope of the invention.

5

10

15

20

25

30

Moreover, nucleic acid molecules encoding ORFX proteins from other species, and thus that have a nucleotide sequence that differs from the human sequence of any of SEQ ID NO:2n-1 (wherein n = 1 to 3161), are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the ORFX cDNAs of the invention can be isolated based on their homology to the human ORFX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

In another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein n = 1 to 3161). In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500 or 750 nucleotides in length. In another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (i.e., nucleic acids encoding ORFX proteins derived from species other than human) or other related sequences (e.g., paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at Tm, 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (e.g., 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions is hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C. This hybridization is followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of any of SEQ ID NO:2n-1 (wherein n = 1 to 3161) corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2n-1 (wherein n=1 to 3161), or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are

hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well known in the art. See, *e.g.*, Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50°C. Other conditions of low stringency that may be used are well known in the art (*e.g.*, as employed for cross-species hybridizations). See, *e.g.*, Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981, *Proc Natl Acad Sci USA* 78: 6789-6792.

Conservative mutations

5

10

15

20

25

30

In addition to naturally-occurring allelic variants of the ORFX sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequence of any of SEQ ID NO:2n-1 (wherein n = 1 to 3161), thereby leading to changes in the amino acid sequence of the encoded ORFX protein, without altering the functional ability of the ORFX protein. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of any of SEQ ID NO:2n-1 (wherein n = 1 to 3161). A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of ORFX without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the ORFX proteins of the present invention, are predicted to be particularly unamenable to alteration.

Amino acid residues that are conserved among members of an ORFX family members are predicted to be less amenable to alteration. For example, an ORFX protein according to the present invention can contain at least one domain (e.g., as shown in Table 1) that is a typically conserved region in an ORFX family member. As such, these conserved domains are not likely to be amenable to mutation. Other amino acid residues, however, (e.g., those that are not conserved or only semi-conserved among members of the ORFX family) may not be as essential for activity and thus are more likely to be amenable to alteration.

5

10

15

20

25

30

Another aspect of the invention pertains to nucleic acid molecules encoding ORFX proteins that contain changes in amino acid residues that are not essential for activity. Such ORFX proteins differ in amino acid sequence from any of any of SEQ ID NO:2n (wherein n = 1 to 3161), yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 75% homologous to the amino acid sequence of any of SEQ ID NO:2n (wherein n = 1 to 3161). Preferably, the protein encoded by the nucleic acid is at least about 80% homologous to any of SEQ ID NO:2n (wherein n = 1 to 3161), more preferably at least about 90%, 95%, 98%, and most preferably at least about 99% homologous to SEQ ID NO:2.

An isolated nucleic acid molecule encoding a ORFX protein homologous to the protein of any of SEQ ID NO:2n (wherein n = 1 to 3161) can be created by introducing one or more nucleotide substitutions, additions or deletions into the corresponding nucleotide sequence, *i.e.* SEQ ID NO:2n-1 for the corresponding n, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NO:2*n*-1 (wherein n = 1 to 3161) by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline,

phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in ORFX is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a ORFX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for ORFX biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:2n-1 (wherein n = 1 to 3161), the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

In one embodiment, a mutant ORFX protein can be assayed for (1) the ability to form protein:protein interactions with other ORFX proteins, other cell-surface proteins, or biologically active portions thereof, (2) complex formation between a mutant ORFX protein and a ORFX receptor; (3) the ability of a mutant ORFX protein to bind to an intracellular target protein or biologically active portion thereof; (e.g., avidin proteins); (4) the ability to bind BRA protein; or (5) the ability to specifically bind an anti-ORFX protein antibody.

Antisense

5

10

15

20

25

30

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:2n-1 (wherein n = 1 to 3161), or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire ORFX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a ORFX protein of any of SEQ ID NO:2n (wherein n = 1 to 3161) or antisense nucleic acids complementary to a ORFX nucleic acid sequence of SEQ ID NO:2n-1 (wherein n = 1 to 3161) are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid

residues (e.g., the protein coding region of a human ORFX that corresponds to any of SEQ ID NO:2n (wherein n = 1 to 3161)). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding ORFX disclosed herein (e.g., SEQ ID NO:2n-1 (wherein n = 1 to 3161)), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of ORFX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of ORFX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of ORFX mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil,

3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a ORFX protein to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

Ribozymes and PNA moieties

5

10

15

20

25

30

Such modifications include, by way of nonlimiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are

carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

5

10

15

20

25

30

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave ORFX mRNA transcripts to thereby inhibit translation of ORFX mRNA. A ribozyme having specificity for a ORFX-encoding nucleic acid can be designed based upon the nucleotide sequence of a ORFX DNA disclosed herein (i.e., SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161)). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a ORFX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, ORFX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

Alternatively, ORFX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the ORFX (e.g., the ORFX promoter and/or enhancers) to form triple helical structures that prevent transcription of the ORFX gene in target cells. See generally, Helene. (1991) Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

In various embodiments, the nucleic acids of ORFX can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide

synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

PNAs of ORFX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of ORFX can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

5

10

15

20

25

30

In another embodiment, PNAs of ORFX can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of ORFX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or

the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

ORFX polypeptides

5

10

15

20

25

The novel protein of the invention includes the ORFX-like protein whose sequence is provided in any of SEQ ID NO:2n (wherein n = 1 to 3161). The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residue shown in FIG. 1 while still encoding a protein that maintains its ORFX-like activities and physiological functions, or a functional fragment thereof. For example, the invention includes the polypeptides encoded by the variant ORFX nucleic acids described above. In the mutant or variant protein, up to 20% or more of the residues may be so changed.

In general, an ORFX -like variant that preserves ORFX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above. Furthermore, without limiting the scope of the invention, positions of any of SEQ ID NO:2n (wherein n = 1 to 3161) may be substitute such that a mutant or variant protein may include one or more substitutions

The invention also includes isolated ORFX proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-ORFX antibodies. In one embodiment, native ORFX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, ORFX proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a ORFX

protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

5

10

15

20

25

30

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the ORFX protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of ORFX protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of ORFX protein having less than about 30% (by dry weight) of non-ORFX protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-ORFX protein, still more preferably less than about 10% of non-ORFX protein, and most preferably less than about 5% non-ORFX protein. When the ORFX protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of ORFX protein in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of ORFX protein having less than about 30% (by dry weight) of chemical precursors or non-ORFX chemicals, more preferably less than about 20% chemical precursors or non-ORFX chemicals, still more preferably less than about 10% chemical precursors or non-ORFX chemicals, and most preferably less than about 5% chemical precursors or non-ORFX chemicals.

Biologically active portions of a ORFX protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the ORFX protein, e.g., the amino acid sequence shown in SEQ ID NO:2 that include fewer amino acids than the full length ORFX proteins, and exhibit at least one activity of a ORFX protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the ORFX protein. A biologically active portion of a ORFX protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length.

A biologically active portion of a ORFX protein of the present invention may contain at least one of the above-identified domains conserved between the FGF family of proteins. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native ORFX protein.

In an embodiment, the ORFX protein has an amino acid sequence shown in any of SEQ ID NO:2n (wherein n = 1 to 3161). In other embodiments, the ORFX protein is substantially homologous to any of SEQ ID NO:2n (wherein n = 1 to 3161) and retains the functional activity of the protein of any of SEQ ID NO:2n (wherein n = 1 to 3161), yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail below. Accordingly, in another embodiment, the ORFX protein is a protein that comprises an amino acid sequence at least about 45% homologous, and more preferably about 55, 65, 70, 75, 80, 85, 90, 95, 98 or even 99% homologous to the amino acid sequence of any of SEQ ID NO:2n (wherein n = 1 to 3161) and retains the functional activity of the ORFX proteins of the corresponding polypeptide having the sequence of SEQ ID NO:2n (wherein n = 1 to 3161).

Determining homology between two or more sequences

5

10

15

20

25

30

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in either of the sequences being compared for optimal alignment between the sequences). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (i.e., as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. See, *Needleman and Wunsch* 1970 *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a

degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NO:2n-1 (wherein n = 1 to 3161).

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (e.g., A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region. The term "percentage of positive residues" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical and conservative amino acid substitutions, as defined above, occur in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of positive residues.

Chimeric and fusion proteins

5

10

15

20

25

30

The invention also provides ORFX chimeric or fusion proteins. As used herein, a ORFX "chimeric protein" or "fusion protein" includes a ORFX polypeptide operatively linked to a non-ORFX polypeptide. A "ORFX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to ORFX, whereas a "non-ORFX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the ORFX protein, e.g., a protein that is different from the ORFX protein and that is derived from the same or a different organism. Within a ORFX fusion protein the ORFX polypeptide can correspond to all or a portion of a ORFX protein. In one embodiment, a ORFX fusion protein comprises at least one biologically active portion of a ORFX protein. In another embodiment, a ORFX fusion protein comprises at least two biologically active portions of a

ORFX protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the ORFX polypeptide and the non-ORFX polypeptide are fused in-frame to each other. The non-ORFX polypeptide can be fused to the N-terminus or C-terminus of the ORFX polypeptide.

5

10

15

20

25

30

For example, in one embodiment a ORFX fusion protein comprises a ORFX polypeptide operably linked to the extracellular domain of a second protein. Such fusion proteins can be further utilized in screening assays for compounds that modulate ORFX activity (such assays are described in detail below).

In another embodiment, the fusion protein is a GST-ORFX fusion protein in which the ORFX sequences are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant ORFX.

In yet another embodiment, the fusion protein is a ORFX protein containing a heterologous signal sequence at its N-terminus. For example, the native ORFX signal sequence can be removed and replaced with a signal sequence from another protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of ORFX can be increased through use of a heterologous signal sequence.

In another embodiment, the fusion protein is a ORFX-immunoglobulin fusion protein in which the ORFX sequences comprising one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The ORFX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ORFX ligand and a ORFX protein on the surface of a cell, to thereby suppress ORFX-mediated signal transduction *in vivo*. In one nonlimiting example, a contemplated ORFX ligand of the invention is an ORFX receptor. The ORFX-immunoglobulin fusion proteins can be used to modulate the bioavailability of a ORFX cognate ligand. Inhibition of the ORFX ligand/ORFX interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (e.g., promoting or inhibiting) cell survival. Moreover, the ORFX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-ORFX antibodies in a subject, to purify ORFX ligands, and in screening assays to identify molecules that inhibit the interaction of ORFX with a ORFX ligand.

A ORFX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A ORFX-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the ORFX protein.

ORFX agonists and antagonists

5

10

15

20

25

30

The present invention also pertains to variants of the ORFX proteins that function as either ORFX agonists (mimetics) or as ORFX antagonists. Variants of the ORFX protein can be generated by mutagenesis, e.g., discrete point mutation or truncation of the ORFX protein. An agonist of the ORFX protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the ORFX protein. An antagonist of the ORFX protein can inhibit one or more of the activities of the naturally occurring form of the ORFX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the ORFX protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the ORFX proteins.

Variants of the ORFX protein that function as either ORFX agonists (mimetics) or as ORFX antagonists can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of the ORFX protein for ORFX protein agonist or antagonist activity. In one

embodiment, a variegated library of ORFX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of ORFX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential ORFX sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of ORFX sequences therein. There are a variety of methods which can be used to produce libraries of potential ORFX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential ORFX sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu Rev Biochem 53:323; Itakura et al. (1984) Science 198:1056; Ike et al. (1983) Nucl Acid Res 11:477.

Polypeptide libraries

5

10

15

20

25

30

In addition, libraries of fragments of the ORFX protein coding sequence can be used to generate a variegated population of ORFX fragments for screening and subsequent selection of variants of a ORFX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a ORFX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the ORFX protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of ORFX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors,

transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recrusive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify ORFX variants (Arkin and Yourvan (1992) PNAS 89:7811-7815; Delgrave *et al.* (1993) Protein Engineering 6:327-331).

Anti-ORFX Antibodies

5

10

15

20

25

30

The invention further encompasses antibodies and antibody fragments, such as F_{ab} or $(F_{ab})_2$ that bind immunospecifically to any of the proteins of the invention.

An isolated ORFX protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind ORFX using standard techniques for polyclonal and monoclonal antibody preparation. Full-length ORFX protein can be used. Alternatively, the invention provides antigenic peptide fragments of ORFX for use as immunogens. The antigenic peptide of ORFX comprises at least 4 amino acid residues of the amino acid sequence shown in any of SEQ ID NO:2n (wherein n = 1 to 3161). The antigenic peptide encompasses an epitope of ORFX such that an antibody raised against the peptide forms a specific immune complex with ORFX. The antigenic peptide may comprise at least 6 aa residues, at least 8 aa residues, at least 10 aa residues, at least 15 aa residues, at least 20 aa residues, or at least 30 aa residues. In one embodiment of the invention, the antigenic peptide comprises a polypeptide comprising at least 6 contiguous amino acids of any of SEQ ID NO:2n (wherein n = 1 to 3161).

In an embodiment of the invention, epitopes encompassed by the antigenic peptide are regions of ORFX that are located on the surface of the protein, *e.g.*, hydrophilic regions. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each incorporated herein by reference in their entirety.

As disclosed herein, an ORFX protein sequence of any of SEQ ID NO:2n (wherein n = 1 to 3161), or derivatives, fragments, analogs or homologs thereof, may be utilized as immunogens in the generation of antibodies that immunospecifically-bind these protein components. The term "antibody" as used herein refers to immunoglobulin molecules and

immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen, such as ORFX. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In a specific embodiment, antibodies to human ORFX proteins are disclosed. Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies to a ORFX protein sequence of any of SEQ ID NO:2n (wherein n = 1 to 3161) or derivative, fragment, analog or homolog thereof. Some of these proteins are discussed below.

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by injection with the native protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly expressed ORFX protein or a chemically synthesized ORFX polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. If desired, the antibody molecules directed against ORFX can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of ORFX. A monoclonal antibody composition thus typically displays a single binding affinity for a particular ORFX protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular ORFX protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see Kohler & Milstein, 1975 *Nature* 256: 495-497); the trioma technique; the human B-cell hybridoma technique (see Kozbor, *et al.*, 1983 *Immunol Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND

CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Each of the above citations are incorporated herein by reference in their entirety

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a ORFX protein (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see *e.g.*, Huse, *et al.*, 1989 *Science* 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a ORFX protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See *e.g.*, U.S. Patent No. 5,225,539. Each of the above citations are incorporated herein by reference. Antibody fragments that contain the idiotypes to a ORFX protein may be produced by techniques known in the art including, but not limited to: (*i*) an $F_{(ab)2}$ fragment produced by pepsin digestion of an antibody molecule; (*ii*) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab)2}$ fragment; (*iii*) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (*iv*) F_v fragments.

Additionally, recombinant anti-ORFX antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Pat. No. 4,816,567; European Patent Application No. 125,023; Better et al.(1988) Science 240:1041-1043; Liu et al. (1987) PNAS 84:3439-3443; Liu et al. (1987) J Immunol. 139:3521-3526; Sun et al. (1987) PNAS 84:214-218; Nishimura et al. (1987) Cancer Res 47:999-1005; Wood et al. (1985) Nature 314:446-449; Shaw et al. (1988), J. Natl Cancer Inst 80:1553-1559); Morrison(1985) Science 229:1202-1207; Oi et al. (1986) BioTechniques 4:214; U.S. Pat. No. 5,225,539; Jones et al. (1986) Nature 321:552-525;

Verhoeyan et al. (1988) Science 239:1534; and Beidler et al. (1988) J Immunol 141:4053-4060. Each of the above citations are incorporated herein by reference.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of a ORFX protein is facilitated by generation of hybridomas that bind to the fragment of a ORFX protein possessing such a domain. Antibodies that are specific for one or more domains within a ORFX protein, e.g., the domain spanning the first fifty amino-terminal residues specific to ORFX when compared to FGF-9, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

5

10

15

20

25

30

Anti-ORFX antibodies may be used in methods known within the art relating to the localization and/or quantitation of a ORFX protein (e.g., for use in measuring levels of the ORFX protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for ORFX proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds [hereinafter "Therapeutics"].

An anti-ORFX antibody (*e.g.*, monoclonal antibody) can be used to isolate ORFX by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-ORFX antibody can facilitate the purification of natural ORFX from cells and of recombinantly produced ORFX expressed in host cells. Moreover, an anti-ORFX antibody can be used to detect ORFX protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the ORFX protein. Anti-ORFX antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or

phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ¹²⁵I, ¹³¹I, ³⁵S or ³H.

ORFX Recombinant Vectors and Host Cells

5

10

15

20

25

30

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding ORFX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to includes promoters, enhancers and other expression control elements

(e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., ORFX proteins, mutant forms of ORFX, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of ORFX in prokaryotic or eukaryotic cells. For example, ORFX can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (1) to increase expression of recombinant protein; (2) to increase the solubility of the recombinant protein; and (3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

5

10

15

20

25

30

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the ORFX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari, *et al.*, (1987) *EMBO J* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (InVitrogen Corp, San Diego, Calif.).

Alternatively, ORFX can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., SF9 cells) include the pAc series (Smith et al. (1983) Mol Cell Biol 3:2156-2165) and the pVL series (Lucklow and Summers (1989) Virology 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J* 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells. See, *e.g.*, Chapters 16 and 17 of Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al. (1987) Genes Dev 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) Adv Immunol 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) EMBO J 8:729-733) and immunoglobulins (Banerji et al. (1983) Cell 33:729-740; Queen and Baltimore (1983) Cell 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) PNAS 86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) Science 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, e.g., the murine hox promoters (Kessel and Gruss (1990) Science 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) Genes Dev 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to ORFX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub *et al.*, "Antisense RNA as a molecular tool for genetic analysis," Reviews--Trends in Genetics, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant

host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

5

10

15

20

25

30

A host cell can be any prokaryotic or eukaryotic cell. For example, ORFX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding ORFX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) ORFX protein. Accordingly, the invention further provides methods for producing ORFX protein using the host cells of the invention. In one embodiment,

the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding ORFX has been introduced) in a suitable medium such that ORFX protein is produced. In another embodiment, the method further comprises isolating ORFX from the medium or the host cell.

Transgenic animals

5

10

15

20

25

30

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which ORFX-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous ORFX sequences have been introduced into their genome or homologous recombinant animals in which endogenous ORFX sequences have been altered. Such animals are useful for studying the function and/or activity of ORFX and for identifying and/or evaluating modulators of ORFX activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous ORFX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing ORFX-encoding nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human ORFX DNA sequence of SEQ ID NO:2n-1 (wherein n = 1 to 3161) can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of the human ORFX gene, such as a mouse ORFX gene, can be isolated based on hybridization to the human ORFX cDNA (described further above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of

expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the ORFX transgene to direct expression of ORFX protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Pat. Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan 1986, In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the ORFX transgene in its genome and/or expression of ORFX mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding ORFX can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a ORFX gene into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the ORFX gene. The ORFX gene can be a human gene (e.g., SEQ ID NO:2n-1 (wherein n=1 to 3161)), but more preferably, is a non-human homologue of a human ORFX gene. For example, a mouse homologue of human ORFX gene of SEQ ID NO:2n-1 (wherein n=1 to 3161) can be used to construct a homologous recombination vector suitable for altering an endogenous ORFX gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous ORFX gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous ORFX gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous ORFX protein). In the homologous recombination vector, the altered portion of the ORFX gene is flanked at its 5' and 3' ends by additional nucleic acid of the ORFX gene to allow for homologous recombination to occur between the exogenous ORFX gene carried by the vector and an endogenous ORFX gene in an embryonic stem cell. The additional flanking ORFX nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector. See e.g., Thomas et al. (1987) Cell 51:503 for a description of

homologous recombination vectors. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced ORFX gene has homologously recombined with the endogenous ORFX gene are selected (see e.g., Li et al. (1992) Cell 69:915).

5

10

15

20

25

30

The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras. See e.g., Bradley 1987, In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) Curr Opin Biotechnol 2:823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, see, e.g., Lakso et al. (1992) PNAS 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of Saccharomyces cerevisiae (O'Gorman et al. (1991) Science 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al.* (1997) *Nature* 385:810-813. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0 phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of

this female foster animal will be a clone of the animal from which the cell, e.g., the somatic cell, is isolated.

Pharmaceutical Compositions

The ORFX nucleic acid molecules, ORFX proteins, and anti-ORFX antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

20

25

30

5

10

15

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor ELTM (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

5

10

15

20

25

30

Sterile injectable solutions can be prepared by incorporating the active compound (e.g., a ORFX protein or anti-ORFX antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use

as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

5

10

15

20

25

30

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Pat. No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by any of a number of routes, e.g., as described in U.S. Patent Nos. 5,703,055. Delivery can thus also include, e.g., intravenous injection, local administration (see U.S. Pat. No. 5,328,470) or stereotactic injection (see e.g., Chen et al. (1994) PNAS 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g., retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

Additional Uses and Methods of the Invention

5

10

15

20

25

30

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: (a) screening assays; (b) detection assays (e.g., chromosomal mapping, cell and tissue typing, forensic biology), (c) predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenomics); and (d) methods of treatment (e.g., therapeutic and prophylactic).

The isolated nucleic acid molecules of the invention can be used to express ORFX protein (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect ORFX mRNA (e.g., in a biological sample) or a genetic lesion in a ORFX gene, and to modulate ORFX activity, as described further below. In addition, the ORFX proteins can be used to screen drugs or compounds that modulate the ORFX activity or expression as well as to treat disorders characterized by insufficient or excessive production of ORFX protein, for

example proliferative or differentiative disorders, or production of ORFX protein forms that have decreased or aberrant activity compared to ORFX wild type protein. In addition, the anti-ORFX antibodies of the invention can be used to detect and isolate ORFX proteins and modulate ORFX activity.

This invention further pertains to novel agents identified by the above described screening assays and uses thereof for treatments as described herein.

Screening Assays

5

10

15

20

25

30

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) that bind to ORFX proteins or have a stimulatory or inhibitory effect on, for example, ORFX expression or ORFX activity.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a ORFX protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) Anticancer Drug Des 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al. (1993) Proc Natl Acad Sci U.S.A. 90:6909; Erb et al. (1994) Proc Natl Acad Sci U.S.A. 91:11422; Zuckermann et al. (1994) J Med Chem 37:2678; Cho et al. (1993) Science 261:1303; Carrell et al. (1994) Angew Chem Int Ed Engl 33:2059; Carell et al. (1994) Angew Chem Int Ed Engl 33:2061; and Gallop et al. (1994) J Med Chem 37:1233.

Libraries of compounds may be presented in solution (e.g., Houghten (1992) Biotechniques 13:412-421), or on beads (Lam (1991) Nature 354:82-84), on chips (Fodor (1993) Nature 364:555-556), bacteria (Ladner U.S. Pat. No. 5,223,409), spores (Ladner USP '409), plasmids (Cull et al. (1992) Proc Natl Acad Sci USA 89:1865-1869) or on phage (Scott and

Smith (1990) Science 249:386-390; Devlin (1990) Science 249:404-406; Cwirla et al. (1990) Proc Natl Acad Sci U.S.A. 87:6378-6382; Felici (1991) J Mol Biol 222:301-310; Ladner above.).

5

10

15

20

25

30

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to a ORFX protein determined. The cell, for example, can of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the ORFX protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the ORFX protein or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ¹²⁵I, ³⁵S, ¹⁴C, or ³H, either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the test compound to preferentially bind to ORFX or a biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the ORFX protein or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of ORFX or a biologically active portion thereof can be accomplished, for example, by determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule. As used herein, a "target molecule" is a molecule with which a ORFX protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a ORFX interacting protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule

associated with the internal surface of a cell membrane or a cytoplasmic molecule. A ORFX target molecule can be a non-ORFX molecule or a ORFX protein or polypeptide of the present invention. In one embodiment, a ORFX target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (e.g., a signal generated by binding of a compound to a membrane-bound ORFX molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with ORFX.

5

10

15

20

25

30

Determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (i.e. intracellular Ca²⁺, diacylglycerol, IP₃, etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising a ORFX-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, e.g., luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a ORFX protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to the ORFX protein or biologically active portion thereof. Binding of the test compound to the ORFX protein can be determined either directly or indirectly as described above. In one embodiment, the assay comprises contacting the ORFX protein or biologically active portion thereof with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the test compound to preferentially bind to ORFX or biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-free assay comprising contacting ORFX protein or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the ORFX protein or

biologically active portion thereof. Determining the ability of the test compound to modulate the activity of ORFX can be accomplished, for example, by determining the ability of the ORFX protein to bind to a ORFX target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of ORFX can be accomplished by determining the ability of the ORFX protein further modulate a ORFX target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

5

10

15

20

25

30

In yet another embodiment, the cell-free assay comprises contacting the ORFX protein or biologically active portion thereof with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the ORFX protein to preferentially bind to or modulate the activity of a ORFX target molecule.

The cell-free assays of the present invention are amenable to use of both the soluble form or the membrane-bound form of ORFX. In the case of cell-free assays comprising the membrane-bound form of ORFX, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of ORFX is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)n, N-dodecyl--N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl)dimethylamminiol-1-propane sulfonate (CHAPSO).

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either ORFX or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to ORFX, or interaction of ORFX with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For

example, GST-ORFX fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or ORFX protein, and the mixture is incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of ORFX binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either ORFX or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated ORFX or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with ORFX or target molecules, but which do not interfere with binding of the ORFX protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or ORFX trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the ORFX or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the ORFX or target molecule.

In another embodiment, modulators of ORFX expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of ORFX mRNA or protein in the cell is determined. The level of expression of ORFX mRNA or protein in the presence of the candidate compound is compared to the level of expression of ORFX mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of ORFX expression based on this comparison. For example, when expression of ORFX mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of ORFX mRNA or protein expression. Alternatively, when expression of ORFX mRNA or protein is less (statistically significantly less) in the presence of the candidate

compound than in its absence, the candidate compound is identified as an inhibitor of ORFX mRNA or protein expression. The level of ORFX mRNA or protein expression in the cells can be determined by methods described herein for detecting ORFX mRNA or protein.

In yet another aspect of the invention, the ORFX proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, e.g., U.S. Pat. No. 5,283,317; Zervos et al. (1993) Cell 72:223-232; Madura et al. (1993) J Biol Chem 268:12046-12054; Bartel et al. (1993) Biotechniques 14:920-924; Iwabuchi et al. (1993) Oncogene 8:1693-1696; and Brent WO94/10300), to identify other proteins that bind to or interact with ORFX ("ORFX-binding proteins" or "ORFX-bp") and modulate ORFX activity. Such ORFX-binding proteins are also likely to be involved in the propagation of signals by the ORFX proteins as, for example, upstream or downstream elements of the ORFX pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for ORFX is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, in vivo, forming a ORFX-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with ORFX.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

Detection Assays

5

10

15

20

25

30

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a

minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample.

The ORFX sequences of the present invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the present invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Pat. No. 5,272,057).

5

10

15

20

25

Furthermore, the sequences of the present invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the ORFX sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The ORFX sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:2n-1 (wherein n=1 to 3161), as described above, can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

Predictive Medicine

5

10

15

20

25

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining ORFX protein and/or nucleic acid expression as well as ORFX activity, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant ORFX expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with ORFX protein, nucleic acid expression or activity. For example, mutations in a ORFX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with ORFX protein, nucleic acid expression or activity.

Another aspect of the invention provides methods for determining ORFX protein, nucleic acid expression or ORFX activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of ORFX in clinical trials.

Use of Partial ORFX Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, *e.g.*, hair or skin, or body fluids, *e.g.*, blood, saliva, or semen

found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, that can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e. another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NOs:__ are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the ORFX sequences or portions thereof, e.g., fragments derived from the noncoding regions of one or more of SEQ ID NO:2n-1 (where n = 1 to 3161), having a length of at least 20 bases, preferably at least 30 bases.

The ORFX sequences described herein can further be used to provide polynucleotide reagents, e.g., labeled or label-able probes that can be used, for example, in an *in situ* hybridization technique, to identify a specific tissue, e.g., brain tissue, etc. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such ORFX probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, e.g., ORFX primers or probes can be used to screen tissue culture for contamination (i.e. screen for the presence of a mixture of different types of cells in a culture).

Predictive Medicine

5

10

15

20

25

30

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining ORFX protein and/or nucleic acid expression as well as ORFX activity, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant ORFX expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with ORFX protein, nucleic

acid expression or activity. For example, mutations in a ORFX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with ORFX protein, nucleic acid expression or activity.

Another aspect of the invention provides methods for determining ORFX protein, nucleic acid expression or ORFX activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of ORFX in clinical trials.

These and other agents are described in further detail in the following sections.

Diagnostic Assays

5

10

15

20

25

30

Other conditions in which proliferation of cells plays a role include tumors, restenosis, psoriasis, Dupuytren's contracture, diabetic complications, Kaposi's sarcoma and rheumatoid arthritis.

An ORFX polypeptide may be used to identify an interacting polypeptide a sample or tissue. The method comprises contacting the sample or tissue with ORFX, allowing formation of a complex between the ORFX polypeptide and the interacting polypeptide, and detecting the complex, if present.

The proteins of the invention may be used to stimulate production of antibodies specifically binding the proteins. Such antibodies may be used in immunodiagnostic procedures to detect the occurrence of the protein in a sample. The proteins of the invention may be used to stimulate cell growth and cell proliferation in conditions in which such growth would be favorable. An example would be to counteract toxic side effects of chemotherapeutic agents on, for example, hematopoiesis and platelet formation, linings of the gastrointestinal tract, and hair follicles. They may also be used to stimulate new cell growth in neurological disorders including, for example, Alzheimer's disease. Alternatively, antagonistic treatments may be administered in which an antibody specifically binding the ORFX -like proteins of the invention

would abrogate the specific growth-inducing effects of the proteins. Such antibodies may be useful, for example, in the treatment of proliferative disorders including various tumors and benign hyperplasias.

Polynucleotides or oligonucleotides corresponding to any one portion of the ORFX nucleic acids of SEQ ID NO:2n-1 (wherein n=1 to 3161) may be used to detect DNA containing a corresponding ORF gene, or detect the expression of a corresponding ORFX gene, or ORFX-like gene. For example, an ORFX nucleic acid expressed in a particular cell or tissue, as noted in Table 2, can be used to identify the presence of that particular cell type.

5

10

15

20

25

30

An exemplary method for detecting the presence or absence of ORFX in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting ORFX protein or nucleic acid (e.g., mRNA, genomic DNA) that encodes ORFX protein such that the presence of ORFX is detected in the biological sample. An agent for detecting ORFX mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to ORFX mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length ORFX nucleic acid, such as the nucleic acid of SEQ ID NO:2n-1 (wherein n=1 to 3161), or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to ORFX mRNA or genomic DNA, as described above. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting ORFX protein is an antibody capable of binding to ORFX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect ORFX mRNA, protein, or genomic DNA in a biological sample in vitro as

well as *in vivo*. For example, *in vitro* techniques for detection of ORFX mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of ORFX protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of ORFX genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of ORFX protein include introducing into a subject a labeled anti-ORFX antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting ORFX protein, mRNA, or genomic DNA, such that the presence of ORFX protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of ORFX protein, mRNA or genomic DNA in the control sample with the presence of ORFX protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of ORFX in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting ORFX protein or mRNA in a biological sample; means for determining the amount of ORFX in the sample; and means for comparing the amount of ORFX in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect ORFX protein or nucleic acid.

Prognostic Assays

5

10

15

20

25

30

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with ORFX protein, nucleic acid expression or activity in, e.g., proliferative or differentiative disorders such as hyperplasias, tumors, restenosis, psoriasis, Dupuytren's

contracture, diabetic complications, or rheumatoid arthritis, etc.; and glia-associated disorders such as cerebral lesions, diabetic neuropathies, cerebral edema, senile dementia, Alzheimer's disease, etc. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant ORFX expression or activity in which a test sample is obtained from a subject and ORFX protein or nucleic acid (e.g., mRNA, genomic DNA) is detected, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (e.g., serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant ORFX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder, such as a proliferative disorder, differentiative disorder, glia-associated disorders, etc. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant ORFX expression or activity in which a test sample is obtained and ORFX protein or nucleic acid is detected (e.g., wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant ORFX expression or activity.)

The methods of the invention can also be used to detect genetic lesions in a ORFX gene, thereby determining if a subject with the lesioned gene is at risk for, or suffers from, a proliferative disorder, differentiative disorder, glia-associated disorder, etc. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding a ORFX-protein, or the mis-expression of the ORFX gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of (1) a deletion of one or more nucleotides from a ORFX gene; (2) an addition of one or more nucleotides to a ORFX gene; (3) a substitution of one or more nucleotides of a ORFX gene, (4) a chromosomal

rearrangement of a ORFX gene; (5) an alteration in the level of a messenger RNA transcript of a ORFX gene, (6) aberrant modification of a ORFX gene, such as of the methylation pattern of the genomic DNA, (7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a ORFX gene, (8) a non-wild type level of a ORFX-protein, (9) allelic loss of a ORFX gene, and (10) inappropriate post-translational modification of a ORFX-protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a ORFX gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) Science 241:1077-1080; and Nakazawa et al. (1994) PNAS 91:360-364), the latter of which can be particularly useful for detecting point mutations in the ORFX-gene (see Abravaya et al. (1995) Nucl Acids Res 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to a ORFX gene under conditions such that hybridization and amplification of the ORFX gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli et al., 1990, Proc Natl Acad Sci USA 87:1874-1878), transcriptional amplification system (Kwoh, et al., 1989, Proc Natl Acad Sci USA 86:1173-1177), Q-Beta Replicase (Lizardi et al, 1988, BioTechnology 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a ORFX gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Pat. No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

5

10

15

20

25

30

In other embodiments, genetic mutations in ORFX can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin et al. (1996) Human Mutation 7: 244-255; Kozal et al. (1996) Nature Medicine 2: 753-759). For example, genetic mutations in ORFX can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin et al. above. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the ORFX gene and detect mutations by comparing the sequence of the sample ORFX with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert (1977) *PNAS* 74:560 or Sanger (1977) *PNAS* 74:5463. It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve et al., (1995) Biotechniques 19:448), including sequencing by mass spectrometry (see, e.g., PCT International Publ. No. WO 94/16101; Cohen et al. (1996) Adv Chromatogr 36:127-162; and Griffin et al. (1993) Appl Biochem Biotechnol 38:147-159).

Other methods for detecting mutations in the ORFX gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA

heteroduplexes (Myers et al. (1985) Science 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type ORFX sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton et al (1988) Proc Natl Acad Sci USA 85:4397; Saleeba et al (1992) Methods Enzymol 217:286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in ORFX cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a ORFX sequence, *e.g.*, a wild-type ORFX sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Pat. No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in ORFX genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita et al. (1989) Proc Natl Acad Sci USA: 86:2766, see also Cotton (1993) Mutat Res 285:125-144; Hayashi (1992) Genet Anal Tech Appl 9:73-79). Single-stranded DNA fragments of sample and control ORFX nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting

alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA, rather than DNA, in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. See, e.g., Keen et al. (1991) Trends Genet 7:5.

5

10

15

20

25

30

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). See, e.g., Myers et al (1985) Nature 313:495. When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA. See, e.g., Rosenbaum and Reissner (1987) Biophys Chem 265:12753.

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found. See, e.g., Saiki et al. (1986) Nature 324:163); Saiki et al. (1989) Proc Natl Acad. Sci USA 86:6230. Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection. See, *e.g.*, Gasparini *et al* (1992) *Mol Cell Probes* 6:1. It is anticipated that in certain embodiments amplification may also be performed using Taq ligase

for amplification. See, e.g., Barany (1991) Proc Natl Acad Sci USA 88:189. In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a ORFX gene.

Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which ORFX is expressed may be utilized in the prognostic assays described herein. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

Pharmacogenomics

5

10

15

20

25

30

Agents, or modulators that have a stimulatory or inhibitory effect on ORFX activity (e.g., ORFX gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (e.g., neurological, cancerrelated or gestational disorders) associated with aberrant ORFX activity. In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (e.g., drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of ORFX protein, expression of ORFX nucleic acid, or mutation content of ORFX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum, 1996, *Clin Exp Pharmacol Physiol*, 23:983-985 and Linder, 1997, *Clin Chem*, 43:254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic

conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

5

10

15

20

25

30

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of ORFX protein, expression of ORFX nucleic acid, or mutation content of ORFX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a ORFX modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring Clinical Efficacy

5

10

15

20

25

30

Monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of ORFX (e.g., the ability to modulate aberrant cell proliferation and/or differentiation) can be applied in basic drug screening and in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase ORFX gene expression, protein levels, or upregulate ORFX activity, can be monitored in clinical trials of subjects exhibiting decreased ORFX gene expression, protein levels, or downregulated ORFX activity.

Alternatively, the effectiveness of an agent determined by a screening assay to decrease ORFX gene expression, protein levels, or downregulate ORFX activity, can be monitored in clinical trials of subjects exhibiting increased ORFX gene expression, protein levels, or upregulated ORFX activity. In such clinical trials, the expression or activity of ORFX and, preferably, other genes that have been implicated in, for example, a proliferative or neurological disorder, can be used as a "read out" or marker of the responsiveness of a particular cell.

For example, genes, including ORFX, that are modulated in cells by treatment with an agent (e.g., compound, drug or small molecule) that modulates ORFX activity (e.g., identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of ORFX and other genes implicated in the disorder. The levels of gene expression (i.e., a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of ORFX or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, protein, peptide, nucleic acid, peptidomimetic, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a ORFX protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more

post-administration samples from the subject; (iv) detecting the level of expression or activity of the ORFX protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the ORFX protein, mRNA, or genomic DNA in the pre-administration sample with the ORFX protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of ORFX to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of ORFX to lower levels than detected, i.e., to decrease the effectiveness of the agent.

Methods of Treatment

5

10

15

20

25

30

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant ORFX expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, (*i*) a ORFX polypeptide, or analogs, derivatives, fragments or homologs thereof; (*ii*) antibodies to a ORFX peptide; (*iii*) nucleic acids encoding a ORFX peptide; (*iv*) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to a ORFX peptide) that are utilized to "knockout" endogenous function of a ORFX peptide by homologous recombination (see, *e.g.*, Capecchi, 1989, *Science* 244: 1288-1292); or (*v*) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between a ORFX peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (i.e., are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized

include, but are not limited to, a ORFX peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (e.g., from biopsy tissue) and assaying it in vitro for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a ORFX peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (e.g., by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (e.g., Northern assays, dot blots, in situ hybridization, etc.).

5

10

15

20

25

30

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant ORFX expression or activity, by administering to the subject an agent that modulates ORFX expression or at least one ORFX activity. Subjects at risk for a disease that is caused or contributed to by aberrant ORFX expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the ORFX aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of ORFX aberrancy, for example, a ORFX agonist or ORFX antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

Another aspect of the invention pertains to methods of modulating ORFX expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of ORFX protein activity associated with the cell. An agent that modulates ORFX protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a ORFX protein, a peptide, a ORFX peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more ORFX protein activity. Examples of such stimulatory agents include active ORFX protein and a nucleic acid molecule encoding ORFX that has been introduced into the cell. In another embodiment, the agent inhibits one or more ORFX protein activity. Examples of such inhibitory agents include antisense ORFX nucleic acid molecules and anti-ORFX antibodies. These modulatory methods can be performed *in vitro*

(e.g., by culturing the cell with the agent) or, alternatively, in vivo (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a ORFX protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) ORFX expression or activity. In another embodiment, the method involves administering a ORFX protein or nucleic acid molecule as therapy to compensate for reduced or aberrant ORFX expression or activity.

Determination of the Biological Effect of a Therapeutic

In various embodiments of the present invention, suitable *in vitro* or *in vivo* assays are utilized to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

In various specific embodiments, *in vitro* assays may be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy may be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly, for *in vivo* testing, any of the animal model system known in the art may be used prior to administration to human subjects.

Malignancies

Some ORFX polypeptides are expressed in cancerous cells (see, e.g., Tables 1 and 2). Accordingly, the corresponding ORF protein is involved in the regulation of cell proliferation. Accordingly, Therapeutics of the present invention may be useful in the therapeutic or prophylactic treatment of diseases or disorders that are associated with cell hyperproliferation and/or loss of control of cell proliferation (e.g., cancers, malignancies and tumors). For a review of such hyperproliferation disorders, see e.g., Fishman, et al., 1985. MEDICINE, 2nd ed., J.B. Lippincott Co., Philadelphia, PA.

Therapeutics of the present invention may be assayed by any method known within the art for efficacy in treating or preventing malignancies and related disorders. Such assays include,

but are not limited to, *in vitro* assays utilizing transformed cells or cells derived from the patient's tumor, as well as *in vivo* assays using animal models of cancer or malignancies. Potentially effective Therapeutics are those that, for example, inhibit the proliferation of tumor-derived or transformed cells in culture or cause a regression of tumors in animal models, in comparison to the controls.

In the practice of the present invention, once a malignancy or cancer has been shown to be amenable to treatment by modulating (*i.e.*, inhibiting, antagonizing or agonizing) activity, that cancer or malignancy may subsequently be treated or prevented by the administration of a Therapeutic that serves to modulate protein function.

Premalignant conditions

5

10

15

20

25

The Therapeutics of the present invention that are effective in the therapeutic or prophylactic treatment of cancer or malignancies may also be administered for the treatment of pre-malignant conditions and/or to prevent the progression of a pre-malignancy to a neoplastic or malignant state. Such prophylactic or therapeutic use is indicated in conditions known or suspected of preceding progression to neoplasia or cancer, in particular, where non-neoplastic cell growth consisting of hyperplasia, metaplasia or, most particularly, dysplasia has occurred. For a review of such abnormal cell growth see *e.g.*, Robbins & Angell, 1976. BASIC PATHOLOGY, 2nd ed., W.B. Saunders Co., Philadelphia, PA.

Hyperplasia is a form of controlled cell proliferation involving an increase in cell number in a tissue or organ, without significant alteration in its structure or function. For example, it has been demonstrated that endometrial hyperplasia often precedes endometrial cancer. Metaplasia is a form of controlled cell growth in which one type of mature or fully differentiated cell substitutes for another type of mature cell. Metaplasia may occur in epithelial or connective tissue cells. Dysplasia is generally considered a precursor of cancer, and is found mainly in the epithelia. Dysplasia is the most disorderly form of non-neoplastic cell growth, and involves a loss in individual cell uniformity and in the architectural orientation of cells. Dysplasia characteristically occurs where there exists chronic irritation or inflammation, and is often found in the cervix, respiratory passages, oral cavity, and gall bladder.

Alternatively, or in addition to the presence of abnormal cell growth characterized as hyperplasia, metaplasia, or dysplasia, the presence of one or more characteristics of a transformed or malignant phenotype displayed either *in vivo* or *in vitro* within a cell sample derived from a patient, is indicative of the desirability of prophylactic/therapeutic administration of a Therapeutic that possesses the ability to modulate activity of An aforementioned protein. Characteristics of a transformed phenotype include, but are not limited to: (*i*) morphological changes; (*ii*) looser substratum attachment; (*iii*) loss of cell-to-cell contact inhibition; (*iv*) loss of anchorage dependence; (*v*) protease release; (*vi*) increased sugar transport; (*vii*) decreased serum requirement; (*viii*) expression of fetal antigens, (*ix*) disappearance of the 250 kDal cell-surface protein, and the like. See *e.g.*, Richards, *et al.*, 1986. MOLECULAR PATHOLOGY, W.B. Saunders Co., Philadelphia, PA.

5

10

15

20

25

In a specific embodiment of the present invention, a patient that exhibits one or more of the following predisposing factors for malignancy is treated by administration of an effective amount of a Therapeutic: (i) a chromosomal translocation associated with a malignancy (e.g., the Philadelphia chromosome (bcr/abl) for chronic myelogenous leukemia and t(14;18) for follicular lymphoma, etc.); (ii) familial polyposis or Gardner's syndrome (possible forerunners of colon cancer); (iii) monoclonal gammopathy of undetermined significance (a possible precursor of multiple myeloma) and (iv) a first degree kinship with persons having a cancer or pre-cancerous disease showing a Mendelian (genetic) inheritance pattern (e.g., familial polyposis of the colon, Gardner's syndrome, hereditary exostosis, polyendocrine adenomatosis, Peutz-Jeghers syndrome, neurofibromatosis of Von Recklinghausen, medullary thyroid carcinoma with amyloid production and pheochromocytoma, retinoblastoma, carotid body tumor, cutaneous melanocarcinoma, intraocular melanocarcinoma, xeroderma pigmentosum, ataxia telangiectasia, Chediak-Higashi syndrome, albinism, Fanconi's aplastic anemia and Bloom's syndrome).

In another embodiment, a Therapeutic of the present invention is administered to a human patient to prevent the progression to breast, colon, lung, pancreatic, or uterine cancer, or melanoma or sarcoma.

Hyperproliferative and dysproliferative disorders

In one embodiment of the present invention, a Therapeutic is administered in the therapeutic or prophylactic treatment of hyperproliferative or benign dysproliferative disorders. The efficacy in treating or preventing hyperproliferative diseases or disorders of a Therapeutic of the present invention may be assayed by any method known within the art. Such assays include in vitro cell proliferation assays, in vitro or in vivo assays using animal models of hyperproliferative diseases or disorders, or the like. Potentially effective Therapeutics may, for example, promote cell proliferation in culture or cause growth or cell proliferation in animal models in comparison to controls.

Specific embodiments of the present invention are directed to the treatment or prevention of cirrhosis of the liver (a condition in which scarring has overtaken normal liver regeneration processes); treatment of keloid (hypertrophic scar) formation causing disfiguring of the skin in which the scarring process interferes with normal renewal; psoriasis (a common skin condition characterized by excessive proliferation of the skin and delay in proper cell fate determination); benign tumors; fibrocystic conditions and tissue hypertrophy (e.g., benign prostatic hypertrophy).

Neurodegenerative disorders

5

10

15

20

25

Some ORFX proteins are found in cell types have been implicated in the deregulation of cellular maturation and apoptosis, which are both characteristic of neurodegenerative disease. Accordingly, Therapeutics of the invention, particularly but not limited to those that modulate (or supply) activity of an aforementioned protein, may be effective in treating or preventing neurodegenerative disease. Therapeutics of the present invention that modulate the activity of an aforementioned protein involved in neurodegenerative disorders can be assayed by any method known in the art for efficacy in treating or preventing such neurodegenerative diseases and disorders. Such assays include *in vitro* assays for regulated cell maturation or inhibition of apoptosis or *in vivo* assays using animal models of neurodegenerative diseases or disorders, or any of the assays described below. Potentially effective Therapeutics, for example but not by way of limitation, promote regulated cell maturation and prevent cell apoptosis in culture, or reduce neurodegeneration in animal models in comparison to controls.

Once a neurodegenerative disease or disorder has been shown to be amenable to treatment by modulation activity, that neurodegenerative disease or disorder can be treated or prevented by administration of a Therapeutic that modulates activity. Such diseases include all degenerative disorders involved with aging, especially osteoarthritis and neurodegenerative disorders.

Disorders related to organ transplantation

Some ORFX can be associated with disorders related to organ transplantation, in particular but not limited to organ rejection. Therapeutics of the invention, particularly those that modulate (or supply) activity, may be effective in treating or preventing diseases or disorders related to organ transplantation. Therapeutics of the invention (particularly Therapeutics that modulate the levels or activity of an aforementioned protein) can be assayed by any method known in the art for efficacy in treating or preventing such diseases and disorders related to organ transplantation. Such assays include *in vitro* assays for using cell culture models as described below, or *in vivo* assays using animal models of diseases and disorders related to organ transplantation, see *e.g.*, below. Potentially effective Therapeutics, for example but not by way of limitation, reduce immune rejection responses in animal models in comparison to controls.

Accordingly, once diseases and disorders related to organ transplantation are shown to be amenable to treatment by modulation of activity, such diseases or disorders can be treated or prevented by administration of a Therapeutic that modulates activity.

20 Cardiovascular Disease

5

10

15

25

GENX has been implicated in cardiovascular disorders, including in atherosclerotic plaque formation. Diseases such as cardiovascular disease, including cerebral thrombosis or hemorrhage, ischemic heart or renal disease, peripheral vascular disease, or thrombosis of other major vessel, and other diseases, including diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage disease, systemic lupus erythematosus, homocysteinemia, and familial protein or lipid processing diseases, and the like, are either directly or indirectly associated with atherosclerosis. Accordingly, Therapeutics of the invention, particularly those that modulate (or supply) activity or formation may be effective in treating or preventing

atherosclerosis-associated diseases or disorders. Therapeutics of the invention (particularly Therapeutics that modulate the levels or activity) can be assayed by any method known in the art, including those described below, for efficacy in treating or preventing such diseases and disorders.

5

10

15

20

25

A vast array of animal and cell culture models exist for processes involved in atherosclerosis. A limited and non-exclusive list of animal models includes knockout mice for premature atherosclerosis (Kurabayashi and Yazaki, 1996, Int. Angiol. 15: 187-194), transgenic mouse models of atherosclerosis (Kappel et al., 1994, FASEB J. 8: 583-592), antisense oligonucleotide treatment of animal models (Callow, 1995, Curr. Opin. Cardiol. 10: 569-576), transgenic rabbit models for atherosclerosis (Taylor, 1997, Ann. N.Y. Acad. Sci 811: 146-152), hypercholesterolemic animal models (Rosenfeld, 1996, Diabetes Res. Clin. Pract. 30 Suppl.: 1-11), hyperlipidemic mice (Paigen et al., 1994, Curr. Opin. Lipidol. 5: 258-264), and inhibition of lipoxygenase in animals (Sigal et al., 1994, Ann. N.Y. Acad. Sci. 714: 211-224). In addition, in vitro cell models include but are not limited to monocytes exposed to low density lipoprotein (Frostegard et al., 1996, Atherosclerosis 121: 93-103), cloned vascular smooth muscle cells (Suttles et al., 1995, Exp. Cell Res. 218: 331-338), endothelial cell-derived chemoattractant exposed T cells (Katz et al., 1994, J. Leukoc. Biol. 55: 567-573), cultured human aortic endothelial cells (Farber et al., 1992, Am. J. Physiol. 262: H1088-1085), and foam cell cultures (Libby et al., 1996, Curr Opin Lipidol 7: 330-335). Potentially effective Therapeutics, for example but not by way of limitation, reduce foam cell formation in cell culture models, or reduce atherosclerotic plaque formation in hypercholesterolemic mouse models of atherosclerosis in comparison to controls.

Accordingly, once an atherosclerosis-associated disease or disorder has been shown to be amenable to treatment by modulation of activity or formation, that disease or disorder can be treated or prevented by administration of a Therapeutic that modulates activity.

Cytokine and Cell Proliferation/Differentiation Activity

A GENX protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered

to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

5

10

15

20

25

30

The activity of a protein of the invention may, among other means, be measured by the following methods: Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by Coligan *et al.*, Greene Publishing Associates and Wiley-Interscience (Chapter 3 and Chapter 7); Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Bertagnoili *et al.*, *J Immunol* 145:1706-1712, 1990; Bertagnolli *et al.*, *Cell Immunol* 133:327-341, 1991; Bertagnolli, *et al.*, *J Immunol* 149:3778-3783, 1992; Bowman *et al.*, *J Immunol* 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described by Kruisbeek and Shevach, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1, pp. 3.12.1-14, John Wiley and Sons, Toronto 1994; and by Schreiber, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan eds. Vol 1 pp. 6.8.1-8, John Wiley and Sons, Toronto 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described by Bottomly *et al.*, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto 1991; deVries *et al.*, *J Exp Med* 173:1205-1211, 1991; Moreau *et al.*, *Nature* 336:690-692, 1988; Greenberger *et al.*, *Proc Natl Acad Sci U.S.A.* 80:2931-2938, 1983; Nordan, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.6.1-5, John Wiley and Sons, Toronto 1991; Smith *et al.*, *Proc Natl Acad Sci U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11-Bennett, *et al.* In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto 1991; Ciarletta, *et al.*, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and

cytokine production) include, without limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds., Greene Publishing Associates and Wiley-Interscience (Chapter 3Chapter 6, Chapter 7); Weinberger *et al.*, *Proc Natl Acad Sci USA* 77:6091-6095, 1980; Weinberger *et al.*, *Eur J Immun* 11:405-411, 1981; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

5

10

15

20

25

30

A GENX protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by vital (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by vital, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania species., malaria species and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response.

The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or energy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon re-exposure to specific antigen in the absence of the tolerizing agent.

5

10

15

20

25

30

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (e.g., B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to energize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the

immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc Natl Acad Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., FUNDAMENTAL IMMUNOLOGY, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

5

10

15

20

25

30

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and auto-antibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of auto-antibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic vital diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the

patient. Another method of enhancing anti-vital immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (e.g., sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected ex vivo with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection in vivo.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II a chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor

specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods: Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan et al., eds. Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 7); Herrmann et al., Proc Natl Acad Sci USA 78:2488-2492, 1981; Herrmann et al., J Immunol 128:1968-1974, 1982; Handa et al., J Immunol 135:1564-1572, 1985; Takai et al., J Immunol 137:3494-3500, 1986; Takai et al., J Immunol 140:508-512, 1988; Herrmann et al., Proc Natl Acad Sci USA 78:2488-2492, 1981; Herrmann et al., J Immunol 128:1968-1974, 1982; Handa et al., J Immunol 135:1564-1572, 1985; Takai et al., J Immunol 137:3494-3500, 1986; Bowman et al., J Virology 61:1992-1998; Takai et al., J Immunol 140:508-512, 1988; Bertagnolli et al., Cell Immunol 133:327-341, 1991; Brown et al., J Immunol 153:3079-3092, 1994.

5

10

15

20

25

30

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, *J Immunol* 144:3028-3033, 1990; and Mond and Brunswick In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, (eds.) Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 7); Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988; Bertagnolli *et al.*, *J Immunol* 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J Immunol 134:536-544, 1995; Inaba et al., J Exp Med 173:549-559, 1991; Macatonia et al., J Immunol 154:5071-5079, 1995; Porgador et al., J Exp Med 182:255-260, 1995; Nair et al., J Virol 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., J Exp Med 169:1255-1264, 1989; Bhardwaj et al., J Clin Investig 94:797-807, 1994; and Inaba et al., J Exp Med 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Res 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, J Immunol 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., Internat J Oncol 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cell Immunol 155: 111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc Nat Acad Sci USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

5

10

15

20

25

30

A GENX protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow

transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson *et al. Cellular Biology* 15:141-151, 1995; Keller *et al.*, *Mol. Cell. Biol.* 13:473-486, 1993; McClanahan *et al.*, *Blood* 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, In: Culture of Hematopoietic Cells. Freshney, et al. (eds.) Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y 1994; Hirayama et al., Proc Natl Acad Sci USA 89:5907-5911, 1992; McNiece and Briddeli, In: Culture of Hematopoietic Cells. Freshney, et al. (eds.) Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Exp Hematol 22:353-359, 1994; Ploemacher, In: Culture of Hematopoietic Cells. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Spoonceret al., In: Culture of Hematopoietic Cells. Freshhey, et al., (eds.) Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Sutherland, In: Culture of Hematopoietic Cells. Freshney, et al., (eds.) Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

Tissue Growth Activity

5

10

15

20

25

30

A GENX protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation

induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

5

10

15

20

25

30

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendonitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a career as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein

may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nervo injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

5

10

15

20

25

30

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, EPIDERMAL WOUND HEALING, pp. 71-112 (Maibach and Rovee, eds.), Year Book Medical

Publishers, Inc., Chicago, as modified by Eaglstein and Menz, *J. Invest. Dermatol* 71:382-84 (1978).

Activin/Inhibin Activity

5

10

15

20

25

30

A GENX protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin a family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-b group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc Natl Acad Sci USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example,

attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Coligan et al., eds. (Chapter 6.12, Measurement of Alpha and Beta Chemokines 6.12.1-6.12.28); Taub et al. J Clin Invest 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al., Eur J Immunol 25: 1744-1748; Gruberet al. J Immunol 152:5860-5867, 1994; Johnston et al., J Immunol 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

5

10

15

20

25

30

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res.

45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

5

10

15

20

25

30

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell—cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: CURRENT PROTOCOLS IN IMMUNOLOGY, Ed by Coligan, et al., Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc Natl Acad Sci USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell—cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute

conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Tumor Inhibition Activity

5

10

15

20

25

30

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting

deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

Neural disorders in general include Parkinson's disease, Alzheimer's disease, Huntington's disease, multiple sclerosis, amyotrophic lateral sclerosis (ALS), peripheral neuropathy, tumors of the nervous system, exposure to neurotoxins, acute brain injury, peripheral nerve trauma or injury, and other neuropathies, epilepsy, and/or tremors.

10 EQUIVALENTS

5

15

20

From the foregoing detailed description of the specific embodiments of the invention, it should be apparent that particular novel compositions and methods involving nucleic acids, polypeptides, antibodies, detection and treatment have been described. Although these particular embodiments have been disclosed herein in detail, this has been done by way of example for purposes of illustration only, and is not intended to be limiting with respect to the scope of the appended claims that follow. In particular, it is contemplated by the inventors that various substitutions, alterations, and modifications may be made as a matter of routine for a person of ordinary skill in the art to the invention without departing from the spirit and scope of the invention as defined by the claims. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

Table 1

ORF#	ORF# Internal Identification	Protein similarity	Protein domain	Protein	Cells or Tissues in which Gene is Expressed
	Number			Classification	
-	13076366 (1, 2)	Novel Protein sim. GBank gi 4691395 emb CAB41562.1 - (AL049727) putative large secreted protein [Streptomyces coelicotor]		UNCLASSIFIED	264636
2	80248091 (3. 4)	Novel Protein sim. GBank gil2829506jspjP71559jSUCC_MYCTU - SUCCINYL-COA SYNTHETASE BETA CHAIN (SCS-BETA)	Contains protein domain (PF00549) - UNCLASSIFIED CoA-ligases		264907, 264600, 264602, 264762, 264769, 264689, 264638, 264567
ლ	80415924 (5, 6)			UNCLASSIFIED	264910, 264604, 264634, 264905, 264636, 264691, 264907, 264692, 264629
	82018837 (7, 8)				264908, 264909, 264760, 264628, 264635
5	79970035 (9, 10)				22279002, 264563
9	79842462 (11, 12)		Contains protein domain (PF00127) - UNCLASSIFIED Copper binding proteins,		264908
			plastocyanin/azurin family		
~	85515576 (13, 14)	Novel Protein sim. GBank gil4415926 gb AAD20157 - (AC006282) unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	20281099, 35696052, 264508, 264509, 264905, 264906, 264906, 264906, 264909, 264909, 264901, 264901, 264901, 264901, 264901, 264909, 264909, 264909, 264909, 264990, 264990, 264990, 264760, 264760, 264683, 264683, 264682, 264683, 3657109, 264628, 264629, 35696423, 264631, 264632, 264634, 264634, 264634, 264634, 264634, 264634, 264634, 264634, 264639, 18108385, 264583, 264564, 264566, 264569, 26469
					264486
ω	56924278 (15, 16)	Novel Protein sim. GBank gi 585562 sp Q06456 NIRB_KLEPN - NITRITE REDUCTASE (NAD(P)H) LARGE SUBUNIT	_		264907
o.	79394457 (17, 18)				265007, 265019, 263972
10	79556459 (19, 20)			UNCLASSIFIED	264906
7	20414027 (21, 22)				264605
5	94141210 (23, 24)	Novel Protein sim. GBank gij3878145 emb CAA99871 - (Z75543) similar to potassium channel protein [Caenorhabditis elegans]		misc_channel	264259, 265007, 83373044
13	20750551 (25, 26)			UNCLASSIFIED	264556, 264557, 264564
4	95105114 (27, 28)	Novel Protein sim. GBank gij2832781 emb CAA12645 - Contains pr (AJ225805) inward potassium channel alpha subunit [Egeria Ank repeat densa]	Contains protein domain (PF00023) - potassium_channel Ank repeat		35696286, 35696052, 264510, 35695917, 264691, 264628, 35696423, 264555, 264558, 264559, 83373044
15	20458307 (29, 30)	Novel Protein sim. GBank gi 1710791 sp Q10234 RT05_SCHPO - PROBABLE MITOCHONDRIAL 40S RIBOSOMAL PROTEIN SS	Contains protein domain (PF00333) - ribosomalprot Ribosomal protein S5		264604
16	20760356 (31, 32)				264555

17	20292744 (33, 34)	Novel Protein sim. GBank gij1174884[spjP44391]URE1_HAEIN - UREASE ALPHA SUBUNIT (UREA AMIDOHYDROLASE)	Contains protein domain (PF00449) - Urease		264600
18	80246804 (35, 36)	Novel Protein sim. GBank gi 2281102 (AC002333) - SF16 isolog [Arabidopsis thaliana]			29331827, 264555, 264557, 264638, 264558
19	80076624 (37, 38)			UNCLASSIFIED	22278996, 264907, 264910, 264600, 264693
20	20724558 (39, 40)	Novel Protein sim. GBank gi[2506112 sp P43672 UUP_ECOLI - ABC TRANSPORTER ATP-BINDING PROTEIN UUP		transport	264602
7	80417554 (41 42)	Novel Protein sim GBank		LINC! ASSIFIED	22278995 264906 265008 265010 265011
<u>.</u>	(3r ,1r) rest 1100	nover Florein sint. Gbarin gij1730203jspjP50442jGATM_RAT - GLYCINE		טואכניאסטורובי	22270335, 204300, 205000, 205010, 205011, 264602, 264605, 264766, 264688, 21906764,
		AMIDINOTRANSFERASE PRECURSOR (L-			264691, 18108376, 264636, 18108387,
		ARGININE:GLYCINE AMIDINOTRANSFERASE)			264486
22	11705858 (43, 44)				264685
R	80419176 (45, 46)	Novel Protein sim. GBank gil 1877329lemblCAB070771 -	Contains protein domain (PF00441) - dehydrogenase	dehydrogenase	264488, 264907, 264909, 264600, 264602.
1		(292771) fadE25 [Mycobacterium tuberculosis]	Acyl-CoA dehydrogenase		264603, 264605, 264682, 264766, 32833986,
					704030, 204400
74	20291697 (47, 48)				264600
- 1	80253774 (49, 50)				264593
	80255394 (51, 52)			UNCLASSIFIED	22278996, 56182435, 265018, 264566
	80235795 (53, 54)	Novel Protein sim. GBank gil4808369 emb CAB42783.1 -	Contains protein domain (PF00253) - ribosomalprot	ribosomalprot	18108370, 35696423, 264635, 264555
		(AL049841) putative 30S ribosomal protein S14 [Streptomyces coelicolor]	Ribosomal protein S14p/S29e		
8 8	79483561 (55, 56)			UNCLASSIFIED	264638
82	82448765 (57, 58)	Novel Protein sim, GBank	Contains protein domain (PF00365) -	kinase	264601, 264762, 264766, 264769, 264636
		gij3122290jspjO08333jK6PF_STRCO - 6- PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE) (ATP-PFK)	Phosphofructokinase		
<u>e</u>	79199333 (59, 60)			UNCLASSIFIED	264908, 265019, 264687, 21906764,
į	10040150 (61 62)				00700612
_	13040130 (01, 02)			UNCLASSIFIED	+20+02
32	82449495 (63, 64)	Novel Protein sim. GBank gi 3560504 (AF027770) - unknown [Mycobacterium smegmatis]		UNCLASSIFIED	264905, 264605, 264762, 264766, 264687, 264689
ဗ္ဗ	79582628 (65, 66)	Novel Protein sim. GBank gi[2129003 pir G64507 - hypothetical protein MJ1665 - Methanococcus jannaschii		UNCLASSIFIED	264687
\$	87467657 (67, 68)			UNCLASSIFIED	60432289, 264600, 264602, 264760, 18108357, 264769, 265020, 264691
35	95005170 (69, 70)	Novel Protein sim. GBank gi[5420387 emb CAB46679.1] - (A.J243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	264600, 264687, 264558, 264639
ဗ္က	19642042 (71, 72)	Novel Protein sim. GBank gij3287739 sp P73538 BIOB_SYNY3 - BIOTIN SYNTHASE (BIOTIN SYNTHETASE)		synthase	264566
37	20369215 (73, 74)	Novel Protein sim. GBank gi[2313134]gb[AAD07126.1] - (AE000527) delta-1-pyrroline-5-carboxylate dehydrogenase [Helicobacter pylori 26695]		dehydrogenase	264603

88	20466334 (75, 76)	Novel Protein sim. GBank gi[3805970]emb[CAA06231] - IAJ004933 periplasmic nitrate reductase, large subunit		reductase	264605
		[Rhodopseudomonas sp.]			
38	94300715 (77, 78)	Novel Protein sim. GBank gi 1929449 (L63543) - endodermin [Xenopus laevis]	Contains protein domain (PF00207) - complement Alpha-2-macroglobulin family		264905, 264906, 264907, 66712502, 264908, 264909, 264909, 264511, 265009, 264910, 55812038, 264758, 265011, 264762, 264682, 264763,
					264764, 264766, 265022, 264693, 264628,
					264631, 264634, 264635, 264555, 264638,
				1	18108381, 264558, 18108385, 264482
40	20635625 (79, 80)			UNCLASSIFIED	264592
14	80023287 (81, 82)	Novel Protein sim. GBank gi]854065[emb]CAA58337] - (X83413) U88 [Human herbesvirus 6]			264591, 35695917
42	20724566 (83, 84)			UNCLASSIFIED	264602
43	20467069 (85, 86)	Novel Protein sim. GBank gi 3820584 (AF086791) -		synthase	264605
:		carbamoyiphosphate synthelase large subunit [Zymomonas mobilis]			
44	13085297 (87, 88)	Novel Protein sim. GBank	Contains protein domain (PF00958) - synthase	synthase	264769, 264636
		IAA_MYCTU - GMP SYNTHASE ING) (GLUTAMINE GMP SYNTHETASE)	GMP synthase C terminal domain		
45	39384711 (89, 90)	Novel Protein sim. GBank gij 1881738 (U89688) - myosin-l binding protein Acan125 (Acanthamoeba castellanii)		UNCLASSIFIED	264769, 264510, 264508
46	95003398 (91, 92)			ngf	264566
4	11698624 (93, 94)			UNCLASSIFIED	264689
8	79407218 (95, 96)				18108385, 264635, 264828
49	21659844 (97, 98)			UNCLASSIFIED	264603
20	80503996 (99, 100)				264508, 264603, 264769, 264689, 264636,
51	80255569 (101, 102)	Novel Protein sim. GBank gij3411177 (AF076240) - MocC		UNCLASSIFIED	264593, 18108387
		[Rhizobium leguminosarum bv. viciae]			
25	79208528 (103, 104)	Novel Protein sim. GBank gij3914992[sp[Q26264 SM41_HEMPU - 41 KD SPICULE MATRIX PROTEIN PRECURSOR (HSM41) (HPSMC)		struct	264634
53	36996970 (105, 106)	Novel Protein sim. GBank gi[3980411 (AC004561) - putative profine-rich protein [Arabidopsis thaliana]		UNCLASSIFIED	264762
25	79570897 (107, 108)			UNCLASSIFIED	264630, 264909, 264766
92	80202703 (109, 110)	Novel Protein sim. GBank gij1633572 (U52064) - Herpesvirus saimiri ORF73 homolog [Kaposi's sarcoma- associated herpes-like virus]			29331824, 264102, 265018, 18108376
26	8758408 (111, 112)	Novel Protein sim. GBank gil4321580 gb AAD15785 - (AF050114) alginate lyase [Pseudomonas sp. W7]			264604
22	11223386 (113, 114)		Contains protein domain (PF00076) - UNCLASSIFIED	UNCLASSIFIED	264557
			RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)		

88	91227506 (115, 116)	Novel Protein sim. GBank ai 5616074 ab AAD45616.1 AF06194 - (AF061943) protate-	Contains protein domain (PF00069) - kinase Eukarvotic protein kinase domain	kinase	56182575, 264259, 60432049, 35696052, 66712502, 264009, 265008, 265010, 285011
_		derived STE20-like kinase PSK [Homo sapiens]			264681, 29148784, 35695917, 60170615,
					264691, 264692, 264693, 18108374,
					35696423, 56182323, 60432113
29	80077371 (117, 118)	Novel Protein sim. GBank	Contains protein domain (PF00953) - transferase	transferase	264600, 264689, 264638
		gil1172920 sp P45830 RFE_MYCLE - PUTATIVE	Glycosyl transferase		
		UNDECAPRENYL-PHOSPHATE ALPHA-N-			
ļ		ACETYLGLUCUSAMINYL I KANSFERASE			
20	12958341 (119, 120)				264689
<u>6</u>	80426806 (121, 122)	Novel Protein sim. GBank gil1710216 (U79260) - unknown		glycoprotein	264766
		Homo sapiens			
29	13504966 (123, 124)				264630
္မ	16474553 (125, 126)			UNCLASSIFIED	265019
<u>\$</u>	20724578 (127, 128)	Novel Protein sim. GBank gil420945 pir A47041 -		UNCLASSIFIED	264602
		transposase homolog (insertion element ISAE1) -			
		Alcaligenes eutrophus			
65	79326308 (129, 130)	Novel Protein sim. GBank	Contains protein domain (PF00224) - kinase		264563
		gij3122312 sp 006134 KPYK_MYCTU - PYRUVATE	Pyruvate kinase		
9	1005 1001 1001	NIMAGE (TR)			The state of the s
8	40654584 (131, 132)	Novel Protein sim. GBank gij3928723jemb[CAA22219]		transport	22278996, 264558
		(ALU34355) putative ABC transporter (Streptomyces			
ļ		coelicator			
29	78952543 (133, 134)	Novel Protein sim. GBank		dehydrogenase	265021
	,	gi[231985 sp[P30234 DHA_MYCTU - ALANINE			
		DEHYDROGENASE (40 KD ANTIGEN)			
89	79817382 (135, 136)				264909
69	79841764 (137, 138)			UNCLASSIFIED	264908
20	79871329 (139, 140)				264906, 264908
-	65897456 (141, 142)			UNCLASSIFIED	264602 265021
2	87734077 (443 144)	Noval Bratein eim CBank allantenschaus Ansoners			2002, 2002
<u>, </u>	(44) (140) (44)	Novel Frotein sim. Gbank gil4413926jgbjAAU20137 -		UNCLASSIFIED	264488, 264905, 264906, 264907, 264908,
					204311, 203000, 204310, 204735, 07100474,
					264682, 264766, 264686, 264689, 35695917,
					265021, 60170615, 264691, 33657023,
_		T			264692, 264693, 264629, 264631, 264639.
l					22279000
2	80025241 (145, 146)			UNCLASSIFIED	60424179, 264508, 264908, 265007, 264603,
;	20277440 (147 440)				254567, 254569, 254592, 18108387
:	203//410 (14/, 146)			UNCLASSIFIED	264605
7.5	11819032 (149, 150)			UNCLASSIFIED	264689
		(AL021767) vacuolar protein sorting (Schizosaccharomyces			
ŕ	100000000000000000000000000000000000000	leamod			
ę	95105303 (151, 152)	Novel Protein sim. GBank gi 4468811 emb CAB38212 - (AL035601) putative protein [Arabidonsis thaliana]		UNCLASSIFIED	83373044, 264906, 264557
77	10144718 /153 1541	Novel Drotein eim Chank gilb 406 810 mbi Chank			204503
	(+01 '001) 01 (++101)	(X83413) U88 [Human herpesvirus 6]		UNCLASSIFIED	264353
78	8758258 (155, 156)			UNCLASSIFIED	264604
				l	

79	94140190 (157, 158)	Novel Protein sim. GBank oil5689453tdhilBAA83010 11 -	Contains protein domain (DE00189)		25505305 222000 20234022 20234034
			יייייייי ביייייייייייייייייייייייייייי		33090200, 22270998, 29331822, 29331824,
			allegion L.L		29331825, 29331827, 264905, 264906,
					264907, 66712502, 264908, 264909, 265008,
					265009, 264910, 60170831, 55812038,
					33109954, 265017, 265018, 264288, 264768,
					56181562, 21906765, 21906769, 29148784,
					265020, 264690, 264691, 264692, 264693.
					60431528, 35696423, 264631, 264632
					264634 264636 264639 83373044 264564
					264566, 264567
2	62314840 (159, 160)			UNCLASSIFIED	264769, 264601, 265006, 264910, 264604,
					264605, 264634, 264635, 264905, 264762,
					264637, 264592, 264628, 264907, 264691,
					264908, 264567, 264909, 264766
5	2046/247 (161, 162)	Novel Protein sim. GBank		reductase	264605
		gij1723442jspjQ10258jYD2A_SCHPO - HYPOTHETICAL			
		69.0 KD PROTEIN C56FB.10 IN CHROMOSOME			
82	16331388 (163, 164)	Novel Protein sim, GBank dil 2895866 (AF045770).		dobudroconoro	193190
		methylmologic comi aldahida dahida ana ana ana ana		uenyarogenase	SOCKON.
		memymatoriate semi-aldenyde denydrogenase [Oryza			
		sativa			
83	(94741180 (165, 166)	Novel Protein sim. GBank gij3402673 (AC004697) -		UNCLASSIFIED	264488, 264508, 264509, 264905, 264908.
		unknown protein [Arabidopsis thaliana]			284900 284811 284801 284802 284804
					204009, 204011, 204081, 204030, 204084,
	-				264595, 264596, 264758, 264603, 264760,
_					264681, 18108351, 264762, 264682, 264764,
	_				264684, 264766, 264686, 264632, 264637,
					264557, 264638, 264639, 18108385, 264566
\$	80355375 (167, 168)	Novel Protein sim. GBank	4	transport	264508, 264906, 264907, 264908, 264909,
		gij1173364jspjP45380jSAT1_RAT - SULFATE ANION			264910, 264760, 264763, 264764, 264766,
_		TRANSPORTER 1 (CANALICULAR SULFATE			284768 264769 35695855 264636 264637
		TRANSPORTER) (SULFATE/CARBONATE ANTIPORTER)			
82	80499600 (169, 170)	Novel Protein sim. GBank gi[2120998 pir][S70682 -		transferase	264605, 264762, 264687, 264769, 18108374,
		glycosytransferase homolog - Bordetella pertussis	-		264636, 264486
8	39559043 (171, 172)	Novel Protein sim. GBank gil3256023lemblCAA17228.11 -			264910
		(AL021897) hypothetical protein Rv1112 [Mycobacterium			
		(uberculosis)			
87	13856808 (173, 174)			UNCLASSIFIED	264093
				1	20103

95344718 (175, 176)	Novel Protein sim. GBank gil559703 dbj BAA07552 - (D38549) ha1025 is new [Homo sapiens]			5264507, 52646365, 18108398, 65274572, 56182575, 56994075, 35696286, 22276997, 22276998, 22278999, 264092, 264093, 264094, 264095, 26331822, 29331824, 56182181, 66714117, 29331825, 29331826, 5043289, 29331827, 29331828, 35696052, 33656970, 264105, 264908, 29331830, 264905, 264908, 29331830,
				66712502, 52644045, 56182435, 265007, 265009, 264910, 60170831, 264592, 60431735, 60433356, 33657402, 264757, 60433438, 55812038, 264758, 21906754, 52646317, 33108954, 52644296, 87168474, 265011, 87168559, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 265018, 264601, 264601, 264601, 265018, 264601, 264601, 264601, 265018, 264601,
				264 765, 326942.23, 7 1905 763, 7 1905 787, 265020, 256021, 265022, 60170615, 256020, 265021, 265022, 60170615, 27486261, 27486264, 33657349, 35695763, 264628, 263972, 18108374, 55810764, 35696423, 55811676, 65274791, 35695855,
				60431850, 264636, 52644332, 56182323, 60170394, 83373044, 18108385, 18108387, 18108388, 56526486, 87168518, 60432113, 22279000, 22279002, 264482, 264564, 264486
Novel gil171	Novel Protein sim. GBank gil 710383 sp P46352 RIPX_BACSU - PROBABLE INTEGRASE/RECOMBINASE RIPX			264600
_		3	UNCLASSIFIED	264760
Nove gil249	Novel Protein sim. GBank gilz499891fsplP76403JYEGQ_ECOLI - PUTATIVE pp.OTEACE IN BAED OCDE INTERCENIC PECION	ıd	protease	265006
Nove (ALD	Novel Protein sim. GBank gi 3367754 emb CAA20079 - Novel Protein sim. GBank gi 3367754 emb CAA20079 - State of the state	<u> </u>	UNCLASSIFIED	264691
Nove E1-E	Novel Protein sim. GBank gi 2895095 (AF011337) - putative E1-E2 ATPase [Mus musculus]	Α	TPase_associated	ATPase_associated 264907, 264908, 264910, 265009, 264605, 264769
Nove gil11	JACSU - GLYCINE BETAINE	Contains protein domain (PF00571) - transport CBS domain	ansport	264906
				264594

Novel Protein sim. GBank gil4583560 dbi BAA76802.1 - (AB023175) KIAA0958 protein [Homo sapiens] Novel Protein sim. GBank gil4583559 emb CAB40388.1 - (AJ005250) Pypothetical protein [Synachocystis sp.] Novel Protein sim. GBank gil1001693 pbi BAA14430 - (D64002) hypothetical protein [Synachocystis sp.] Novel Protein sim. GBank gil204166 emb CAA11773.1 - FACTOR G (EF-0) Novel Protein sim. GBank gil294166 emb CAA11773.1 - FACTOR G (EF-0) Novel Protein sim. GBank gil294166 emb CAA11773.1 - FACTOR G (EF-0) Novel Protein sim. GBank gil294166 emb CAA13773.1 - FACTOR G (EF-0) Novel Protein sim. GBank gil284166 emb CA5137 - GA1313 Us8 [Human herpesvirus 6] Novel Protein sim. GBank gil28456 emb CA58317 - (K83413) Us8 [Human herpesvirus 6] Novel Protein sim. GBank gil28456 emb CA58317 - (CA39120) similar to AG Canapane (ATP-binding protein) Novel Protein sim. GBank gil233771 emp CA811285 - (CA9120) similar to AG Canapane (ATP-binding protein) Novel Protein sim. GBank gil233771 emp CA811285 - (CA9120) similar to AG Canapane (ATP-binding protein) Reaching subtilis Novel Protein sim. GBank gil233771 emp CA811285 - (CA9120) similar to AG Canapane (ATP-binding protein) Reaching subtilis Novel Protein sim. GBank gil233771 emp CA811285 - (CA9120) similar to AG Canapane (ATP-binding protein) Reaching subtilis Novel Protein sim. GBank gil233771 emp CA811285 - (CA9120) similar to AG Canapane (ATP-binding protein) Reaching subtilis Novel Protein sim. GBank gil233771 emp CA811285 - (CA9120) similar to AG Canapane (ATP-binding protein) Reaching subtilis Reaching subti	264687 264564 UNCLASSIFIED 264908	264687	UNCLASSIFIED 264758, 264603, 264630, 264636, 264637			Z3J310Z4, Z04331, Z19U0734, Z03U19	ATPass associated 20331824 264501 21006751 265010	CLASSIFIED 264511, 265009	nplement 264508	ED		case 264605	synthase 264605		UNCLASSIFIED 265020, 264102, 263972	Τ	UNCLASSIFIED 264605	UNCLASSIFIED 264909	UNCLASSIFIED 264508	
Novel Protein sim. GBank gil4589560 dbi BAA76802.1 - (AB023175) KIAA0958 protein [Homo sapiens] Novel Protein sim. GBank gil4583559 emb CAB40388.1 - (AJ005255) OxyR [Erwinia chrysanthem] Novel Protein sim. GBank gil4583559 emb CAB40388.1 - (BO4002) hypothetical protein [Synechocystis sp.] Novel Protein sim. GBank gil480897 pir [S37485 - gene msg1 protein sim. GBank gil480897 pir [S37485 - gene msg1 protein sim. GBank gil2894166 emb CAA11773.1 - (AJ223998) PCZA361.18 [Amycolatopsis orientalis] Novel Protein sim. GBank gil2894166 emb CAA11773.1 - (AJ223998) PCZA361.18 [Amycolatopsis orientalis] Novel Protein sim. GBank gil2894166 emb CAA58337 - (X83413) U88 [Human herpesvirus 6] Novel Protein sim. GBank gil285417 (U70770) - furrowed Drosophila melanogaster] Novel Protein sim. GBank gil283577 lemb CA5284 - (Z99120) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] Novel Protein sim. GBank gil2330791 emb CAB11255 - (Z99120) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] Novel Protein sim. GBank gil2330791 emb CAB11255 - (Z98601) carboxypeptidase s precursor [Schizosaccharomyces pombe]	NO		חא		Contains protein domain (PF00005) - tran			Contains protein domain (PF00627) - UNI UBA domain	Contains protein domain (PF00084) - con Sushi domain (SCR repeat)	NO.	Helicases conserved C-terminal domain	Contains protein domain (PF00271) - heli	uks		NO.		NO	20	<u>Š</u>	NO
94322125 (191, 192) 79605200 (193, 194) 79427000 (195, 196) 2046524 (197, 198) 79640113 (199, 200) 80203298 (201, 202) 20467258 (203, 204) 20465268 (205, 206) 79605206 (209, 210) 28382058 (211, 212) 80057791 (213, 214) 80237936 (215, 216) 95194148 (217, 218) 79582823 (219, 220) 39585458 (221, 222)			Novel Protein sim. GBank gi 2330791 emb CAB11265 - (298601) carboxypeptidase s precursor Schizosaccharomyces pombe	٦		gij4887229[gb AAD32244.1]AF15075 - (AF150755) microtubule-actin crosslinking factor [Mus musculus]		4 MOUSE - BS4 PROTEIN			SAL		Novel Protein sim. GBank gi 2894166 emb CAA11773.1 - (AJ223998) PCZA361 18 (Amycolatonsis orientalis)	msg1 protein - mouse	Novel Protein sim. GBank gil480897 pir S37485 - gene		Novel Protein sim. GBank gil1169479 sp P43925 EFG_HAEIN - ELONGATION FACTOR G (EF-G)	Novel Protein sim. GBank gij 1001693 jdbj jBAA 10430 j - (D64002) hypothetical protein [Synechocystis sp.]	Novel Protein sim. GBank giļ4583559jemb CAB40388.1 - [AJ005255] OxyR [Erwinia chrysanthemi]	Novel Protein sim. GBank gil4589560jdbj BAA76802.1 - (AB023175) KIAA0958 protein [Homo sapiens]
	79582823 (219, 220) 39565458 (221, 222) 79856038 (223, 224)	79582823 (219, 220)	95194148 (217, 218)		80237936 (215, 216)	(111)	80057791 (213, 214)	28382058 (211, 212)	79605206 (209, 210)	80247572 (207, 208)		20466368 (205, 206)	20467259 (203, 204)		80203298 (201, 202)	79640113 (199, 200)	20466524 (197, 198)	79427000 (195, 196)	79605200 (193, 194)	94322125 (191, 192)

1 1	[Novel Protein sim. GBank gi[2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcomaassociated herpesyirus]		UNCLASSIFIED	52645156, 52645080, 33656970, 264592. 21906754, 27486264, 18108379, 35696423. 264635, 52644332, 18108382. 264635, 5264407, 264000
116	81296689 (231, 232)				264305, 264300, 204301, 204301, 204301, 264758, 264010, 264761, 264685, 264682, 264766, 264685, 264686, 264686, 264682, 264682, 264631, 264632, 264634, 264637, 264638, 264639, 56526486, 264656, 264666, 264666, 264666, 264666, 264666, 264666, 264666, 264666, 264666, 264666, 264666, 264666, 264666, 264666, 264666, 264666, 2646
1117	79636695 (233, 234)				264639, 264693
118	80222170 (235, 236)		Contains protein domain (PF00170) - UNCLASSIFIED bZIP transcription factor	UNCLASSIFIED	263974
119	91013071 (237, 238)	Novel Protein sim. GBank gij732526 (U22327) - alpha2(IV)		UNCLASSIFIED	22278996, 29331824, 60432289, 265007, 60433438, 264603, 264605, 18108351.
		collagen [Caenomabdits elegans]			264769, 264689, 265020, 264534, 27486261, 264558, 83373044, 18108385, 264564
120	8756491 (239, 240)	Novel Protein sim. GBank gi[2131219]pir[]S50157 - cyclindependent kinase chain SRB10 - yeast (Saccharomyces cerevisiae)		kinase	264603
12	80026153 (241, 242)				264595
122	20457620 (243, 244)	Novel Protein sim. GBank gi[2052147 emb CAB08137 - (Z94752) ksgA [Mycobacterium tuberculosis]	Contains protein domain (PF00398) - transferase Ribosomal RNA adenine dimethylases	transferase	264605
123	R758278 (245, 246)				264604
22/2	79104017 (247, 248)	Novel Protein sim GBank		synthase	18108394, 18108397, 265006, 265007,
<u> </u>	(542 : 143) (1540 61	ST_SORBI - GRANULE-BOUND YNTHASE PRECURSOR			265008, 265010, 265011, 18108355, 18108379, 18108380, 18108384
125	87797986 (249, 250)	amate	Contains protein domain (PF00060) - misc_channel Ligand-gated ion channel	misc_channel	264508, 264906, 265009, 264596, 22279002
126	56701283 (251, 252)	Novel Protein sim. GBank gi 5102785 emb CAB45200.1 - (AL079308) putative transcriptional regulator [Streptomyces coelicolor]			264511
127	20467267 (253, 254)			UNCLASSIFIED	264605
128	80248473 (255, 256)	Novel Protein sim. GBank	Contains protein domain (PF00072) - phosphatase	phosphatase	264907, 264909, 264910, 264600, 264601.
		gi 130120 sp P23620 PHOB_PSEAE - PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY PROTEIN PHOB	Response regulator receiver domain		264603, 264605, 18108351, 264693, 264557
129	95290543 (257, 258)	Novel Protein sim. GBank	Contains protein domain (PF00270) - UNCLASSIFIED	UNCLASSIFIED	35696423, 35695855, 264600, 264602,
		gil2606493jspjP38036jYGCB_ECOLI - HYPOTHETICAL 100.5 KD PROTEIN IN IAP-CYSH INTERGENIC REGION	DEAD/DEAH box helicase		264503, 264604, 264605, 264508, 264906, 264564, 264628, 264682, 264565, 264683
1 30	80085583 (259, 260)	Novel Protein sim. GBank gil854065[emb CAA58337] - (X83413) U88 [Human herpesvirus 6]			264634
131	94995022 (261, 262)	Novel Protein sim. GBank gi 1076038 pir 954860 - ABC transporter PstC-2 chain - Mycobacterium tuberculosis	Contains protein domain (PF00528) - Iransport Binding-protein-dependent transport systems inner membrane component	Iransport	18108376, 264769, 29331826, 264689, 22278996, 265021, 264600, 264511, 264601, 264602, 264605, 264905, 264636

435	10087602 (26.1)				
1 5	04630893 (265, 264)	Manal Description Co. of Co. o			264636
3	94030665 (203, 200)	Nover Frotein sint. Gbank gij lov 7 Avljemoj Abd 7 Jubisj - (292771) accA3 [Mycobacterium tuberculosis]	Contains protein domain (P1-00289) - carboxylase Carbamoyl-phosphate synthase (CPSase)		264905, 264689
<u>¥</u>	79834660 (267, 268)	Novel Protein sim. GBank gil4585838 emb CAB40932.1 - (AL049630) putative NADH dehydrogenase [Streptomyces coelicolor]		dehydrogenase	264905, 264605, 265021
135	19885057 (269, 270)	Novel Protein sim. GBank gi 1460074 emb CAB01049 - (Z77250) hypothetical protein Rv2566 [Mycobacterium tuberculosis]			264634
136	79846083 (271, 272)	Novel Protein sim. GBank gij2125896 emb CAA73511 - (Y13070) folylpolyglutamate synthase [Streptomyces coellcolor]		synthase	264508
137	79619770 (273, 274)				264683, 264685, 264686, 264691, 264692, 264693
138	79635971 (275, 276)	Novel Protein sim. GBank gil5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	18108374, 18108385, 33657182, 265010, 22278998, 265006, 265007, 265008, 265009, 264693
139	86688076 (277, 278)	Novel Protein sim. GBank gi 5689912 emb CAB52075.1 - (AL109732) putative mutase [Streptomyces coelicolor A3(2)]	Contains protein domain (PF01817) - dehydrogenase Chorismate mutase		22278996, 265007, 264910, 60433356, 265010, 264602, 264665, 264766, 264688, 264769, 264693, 32833986, 18108374, 18108387
140	79825759 (279, 280)			UNCLASSIFIED	264908
1	20700094 (281, 282)				264600
45	80028104 (283, 284)	Novel Protein sim. GBank gi 3581916 emb CAA20856 - (AL031545) mutS family DNA mismatch repair protein [Schizosaccharomyces pombe]		nuclease	264602, 265017
<u>5</u>	11072274 (285, 286)			UNCLASSIFIED	264600
4	95009102 (287, 288)	Novel Protein sim. GBank gi]3334127 sp P97303 BAC2_MOUSE - TRANSCRIPTION REGULATOR PROTEIN BACH2 (BTB AND CNC HOMOLOG 2)			263978, 264600, 264910, 264532, 264508, 264563, 264569, 264569, 264569, 264569, 264629, 264639
145	80027058 (289, 290)	Novel Protein sim. GBank gij3757569 emb CAA21315 - (AL031863) 1-evidence=predicted by content; 1-method=genefinder;084; 1-method_score=66.31; 1-evidence_end [Drosophila melanogaster]		UNCLASSIFIED	22278996, 264602
146	13085662 (291, 292)	Novel Protein sim. GBank gil140807jspjP24536jY121_BURCE - INSERTION ELEMENT IS402 HYPOTHETICAL 24 KD PROTEIN	Contains protein domain (PF01675) - Transposase		264687
147	94320366 (293, 294)	Novel Protein sim. GBank gi 2827608 emb CAA16663 - (AL021646) uvrD2 [Mycobacterium tuberculosis]		helicase	264905, 264906, 264909, 264510, 265009, 60433356, 264600, 284601, 264604, 264605, 264687, 264769, 18108365, 65274791, 18108387
148	80248804 (295, 296)	Novel Protein sim. GBank gi 2916947 emb CA417585 - (AL021999) hypothetical protein Rv0966 [Mycobacterium tuberculosis]		transport	265009, 265010, 264600, 264602, 264603, 264604, 264605, 264693, 33657109, 264636

149	1802 49173 (297 208)	Novel Drotein cim Chank	1 00000 P		
<u>-</u>	(007 '107) 0 1001-00	gij1723073 spjQ11040 Y081 MYCTU - HYPOTHETICAL	Contains protein domain (Proudus) - transport ABC transporter	transport	265010, 264600, 264601, 264603, 264604, 27486265, 264636
		ABC TRANSPORTER ATP-BINDING PROTEIN CY50.01			
150	20294748 (299, 300)	Novel Protein sim. GBank gij3724125 emb CAA11905 -			264600
		(AJ224340) maltosephosphorylase (Lactobacillus			
į	20706700 2007	Samiancisco			
2	407 40380 (301, 302)	Novel Protein Sim. Gbank	Contains protein domain (PF01676) - UNCLASSIFIED	UNCLASSIFIED	264602
		PHONDHOPENTOWN ITANE	Metanoenzyme superlamny		
		(PHOSPHODEOXYRIBOMUTASE)			
152	95002877 (303, 304)	Novel Protein sim. GBank		peptidase	264602
		gil2497952(spIP55667)Y4TM RHISN - HYPOTHETICAL			
		HYDROLASE/PEPTIDASE Y4TM			
153	80256665 (305, 306)	Novel Protein sim. GBank		UNCLASSIFIED	264593
		gij3123021/spfQ90508/VIT1_FUNHE - VITELLOGENIN I			
		PRECURSOR (VTG I) (CONTAINS: LIPOVITELLIN 1 (LV1);			
		PHOSVITIN (PV); LIPOVITELLIN 2 (LV2))			
154	82305966 (307, 308)				264910, 264762, 264691, 264634
155	20429859 (309, 310)	Novel Protein sim. GBank gil419697 pir JN0443 -	Contains protein domain (PF00140) -	rnapolymerase	264605
		transcription initiation factor sigma homolog hrdB -	Sigma-70 factor		
		Streptomyces aureofaciens			
28	39564742 (311, 312)	Novel Protein sim. GBank gi[628710 pir][S41739 -		UNCLASSIFIED	264565
		hypothetical protein - Escherichia coli			
157	10358887 (313, 314)	Novel Protein sim. GBank gij3695013 (AF052586) - CtrA	Contains protein domain (PF00142) - hydrolase	hydrolase	264691
		[Pseudomonas aeruginosa]	4Fe-4S iron sulfur cluster binding		
			proteins, NifH/frxC family		
128	79761936 (315, 316)	Novel Protein sim. GBank gil1073072 pir C55543 - cmaU		UNCLASSIFIED	264905
		protein - Pseudomonas syringae pv. syringae			
22	78890376 (317, 318)			UNCLASSIFIED	265008
160	11075119 (319, 320)		Contains protein domain (PF00400) -		264605
			WD domain, G-beta repeat		
<u>16</u>	80055007 (321, 322)	Novel Protein sim. GBank	Contains protein domain (PF00327) - ribosomalprot	ribosomalprot	22278996, 264600, 264603, 35695917,
		9i 1173023 sp P46789 RL30_STRCO - 50S RIBOSOMAL	Ribosomal protein L30p/L7e		32833986, 35696423, 264636
163	00016371 (202 204)	FINAL FINAL SOCIETY CONTRACTOR CO			
<u> </u>	9001037 ((323, 324)	Nover Protein sim. Casank gilosu4869[emp[CAB46028.1] - (AL031685) dJ963K23.2 (novel protein) [Homo sapiens]	Contains protein domain (PF00097) - interleukin Zinc finger, C3HC4 type (RING	interleukin	264112, 264532, 22279002
- 1			finger)		
- 1	11692306 (325, 326)			UNCLASSIFIED	264639
- 1	80077902 (327, 328)			UNCLASSIFIED	264905, 264907, 264600
165	10856067 (329, 330)				264691
	88095003 (331, 332)	Novel Protein sim. GBank gil26616911embtCAA157951 -		INC. ASSIFIED	264605 264486
		(AL009204) putalive protease [Streptomyces coelicolor]			200000
167	[16395460 (333, 334)	Novel Protein sim. GBank gil4416478 gb AAD20378 -		UNCLASSIFIED	265010
		(AF125999) transposase [Mycobacterium avium]			
<u>8</u>	80079362 (335, 336)	Novel Protein sim. GBank gij76177[pir][QQECFT - hunothetical 38 8K protein ffet 5' region). Each michig sell			264600
169	80239581 (337, 338)	היקרטווכווכמן סיפון נומן כן ופקוטון - בפטומווטוום כטוו			DEAFER DEAFET DEAFER DEAFER
					204030, 204007, 204000, 204009

420	70612364 (330 340)				
	75012304 (335, 340)				264906
Ξ	95293073 (341, 342)	Novel Protein Sim. GBank g 140888 sp P27847 YIGK_ECOLI - HYPOTHETICAL 15.4 KD PROTEIN IN RECQ-PLDB INTERGENIC REGION [F138]	Contains protein domain (PF01810) - LysE type translocator		264595, 264604
172	37797007 (343, 344)	Novel Protein sim. GBank gij4210905jgb AAD12048.1 - (AF045609) AgIG [Sinorhizoblum meliloti]	Contains protein domain (PF00528) - transport Binding-protein-dependent transport systems inner membrane component	transport	264769
173	57529660 (345, 346)	Novel Protein sim. GBank gij132854 sp P02387 RL2_ECOLI - 50S RIBOSOMAL PROTEIN L2	Bank gi 132854 sp P02387 RL2_ECOL Contains protein domain (PF00181) - irbosomalprot RIbosomal Proteins L2	ribosomalprot	264769
174	95293078 (347, 348)	Novel Protein sim. GBank gil1881350 dbj BAA19377 - (AB001488) PROBABLE TRANSPORT PROTEIN, SIMILAR TO ANTIBIOTIC TRANSPORT-ASSOCIATED PROTEIN ACTII IN STREPTOMYCES COELICOLOR. [Bacillus subtilis]		transport	264510, 264593, 264602, 264603, 264605, 264762, 264693
175	79756270 (349, 350)	Novel Protein sim. GBank gij2072722jemb CAB08326j - (295121) manA [Mycobacterium tuberculosis]		isomerase	264565
176	80066896 (351, 352)	Novel Protein sim. GBank gi[1055198 (U40187) - similar to PIR:A41724 chicken LD (limb deformity) gene product and to formin; also P-rich region similar to collagen [Caenorhabditis elegans]		UNCLASSIFIED	264907, 264910, 264681, 264558
177	86684852 (353, 354)	Novel Protein sim. GBank gi[2326738 emb CAB10952 - [C98268] hypothetical protein Rv1695 [Mycobacterium tuberculosis]	Contains protein domain (PF01513) - UNCLASSIFIED Domain of unknown function		264768, 60424179, 264687, 264688, 264769, 29331826, 60432289, 18108376, 264689, 18108376, 228500, 265020, 284600, 264601, 264601, 264603, 264603, 264604, 264615, 264636, 264636, 264806, 264664, 264564, 264637, 264638, 264486, 60433356, 264766
178	79559526 (355, 356)	Novel Protein sim. GBank gil 1906596 (U81788) - kinesin-73 [Drosophila melanogaster]		struct	264693, 33657109, 264635
179	20263112 (357, 358)			UNCLASSIFIED	264563
180	80488958 (359, 360)	Novel Protein sim. GBank gi 1169367 sp P45256 DNAB_HAEIN - REPLICATIVE DNA HELICASE		l	264769
181	79585369 (361, 362)	Novel Protein sim. GBank gij3170615 (AF059485) - DOC4 [Mus musculus]		UNCLASSIFIED	21906767, 264635, 264639, 18108384
182	80577899 (363, 364)				264259, 35696052, 56182435, 264511, 265018, 33657109, 264555, 264566
183	11614017 (365, 366)	Novel Protein sim. GBank gi 1076627 pir S54172 - inorganic pyrophosphatase (EC 3.6.1.1) - common tobacco		UNCLASSIFIED	264690
<u>\$</u>	10174167 (367, 368)	Novel Protein sim. GBank gil4371280 gb AAD18138 - (AC006260) hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	264510

185 21660822 (369, 370)	(0/	Novel Protein sim. GBank gij3006178jembjCA418398.1j		UNCLASSIFIED	264604
		[Schizosaccharomyces pombe]			
80070329 (371, 372)	, 372)	Novel Protein sim. GBank gil2829802 sp P94408 YCLF_BACSU - HYPOTHETICAL 53.3 KD PROTEIN IN SFP-GERKA INTERGENIC REGION		Iransport	264595
80186611 (373, 374)	374)			UNCLASSIFIED	264369
20464942 (375, 376)	, 376)	Novel Protein sim. GBank gi[3150260]emb[CAA19179] - [AL023634] cyclin [Schizosaccharomyces pombe]		kinase	264605
82338215 (377, 378)	, 378)	Novel Protein sim. GBank gi 2145853 pir 572938 - hflX protein - Mycobaclerium leprae		UNCLASSIFIED	35696052, 264602, 264605, 264762, 264689, 35695917, 18108370, 18108372, 264638,
80086821 (379, 380)	, 380)	Novel Protein sim. GBank gil 1881244 dbj BAA19271 - (AB001488) SIMILAR TO PYRUVATE OXIDASE AND ACETOLACTATE SYNTHASE. [Bacillus subtilis]	Contains protein domain (PF00205) - synthase Thiamine pyrophosphate enzymes	synthase	264563
88095012 (381, 382)	1, 382)	Novel Protein sim. GBank gil120226isp p28725 FKBP_STRCH - FK506-BINDING PROTEIN (PEPTIDYL-PROLYL CIS-TRANS ISOMERASE)	Contains protein domain (PF00254) - isomerase FKBP-type peptidyl-prolyl cis-trans isomerases	isomerase	264508, 264604, 264605, 264769, 264555
		(PPIASE) (ROTAMASE)			
16333379 (383, 384)	, 384)				264567
9910127 (38:	386)				264908, 264693
20464949 (387, 388)	388)				264605
13518389 (389, 390)	(390)	Novel Protein sim. GBank giļ4980892[gb AAD35474.1[AE00171 - (AE001718) ABC <u>transporter, ATP-binding prote</u> in [Thermotoga maritima]		transport	264636
95005569 (391, 392)	, 392)	Novel Protein sim. GBank gil1705461 splP53656 BIOA_ERWHE - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7.8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)	Contains protein domain (PF00202) - gaba Aminotransferases class-III pyridoxal phosphate	gaba	264600, 264689, 264638
80248665 (393, 394)	3, 394)	Novel Protein sim. GBank gi[3122305[sp]Q27778[k6PF_SCHMA - 6- PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE)	Contains protein domain (PF00365) - kinase Phosphofructokinase	kinase	264602, 264682, 264692, 18108374
79163635 (395, 396)	, 396)				264636
78890715 (39	7, 398)	<u>-</u>	Contains protein domain (PF00393) - 6-phosphogluconate dehydrogenases		265008
79413849 (399, 400)	, 400)	Novel Protein sim. GBank gi 2642222 (AF030885) - telomere-associated recQ-like helicase [Ustilago maydis]		UNCLASSIFIED	264595, 264596
00942924 (401, 402)	402)	Novel Prolein sim. GBank gi 2894379 emb CAA74911.1 - (Y14573) ring finger protein [Hordeum vulgare]		UNCLASSIFIED	29331826, 265007, 264512, 33657402, 264596, 265017, 18108351, 264682, 264683, 264767, 264629, 55810764, 264634, 264635, 56182323, 60432113, 22279000

203	705 RROAG (ARTS ARA)	Naval Brakes aim Obest			
	(40), 404)	NOVEL FLOGEN SINI. GEANN gilcantal STAZISIPISTASSEGUEST_USTMA - CHITIN SYNTHASE 1 (CHITIN-UDP ACETYL-GLUCOSAMINYL TRANSFERASE 1)	Contains protein domain (PF01644) - synthase	synthase	264600
203	79843927 (405, 406)	Novel Protein sim. GBank gil1504042 dbj BAA13220 -			22278995 29331822 29331825 29331827
		(D86984) similar to yeast adenylate cyclase (S56776)			264906, 21906754, 264683, 21906766,
204	79855186 (407 408)				21906769, 35696423, 264556
205	10000583 (400 440)	No. of Details and Observe Constitution		UNCLASSIFIED	264909
3	10030303 (403, 410)	Novel Protein sim. GBank gil2633808 emb CAB13310 - (299111) similar to hypothetical proteins [Bacillus subtilis]		transport	264909
506	8758473 (411, 412)			UNCLASSIFIED	264604
207	20754522 (413, 414)	Novel Protein sim. GBank gi 2134381 pir S60678 -		UNCLASSIFIED	264556
		polybromo 1 protein - chicken			
208	20289261 (415, 416)				264605
508	80071069 (417, 418)	Novel Protein sim. GBank			264605, 264689
		gi 2501040 sp 005814 SYP_MYCTU - PROLYL-TRNA SYNTHETASE (PROLINETRNA LIGASE) (PRORS)			
210	80168800 (419, 420)				264905, 264907, 264909, 264766, 264687.
211	80034539 (421, 422)				204031, 204029, 16106374, 264638
ç	0011747474	Married Parkets			263978
<u> </u>	02442474 (423, 424)	Novel Protein Sim. GBank ail5031809frefind 005536 1falSt R., immunoofabrulin		UNCLASSIFIED	264508, 264905, 264906, 264907, 264908,
		Superfamily containing leucine-rich repeat			20400U, 204/02, 204334, 204832, 204834, 264635, 264639, 264486
213	80249562 (425, 426)	Novel Protein sim. GBank	Contains protein domain (PF00330) - Isomerase	isomerase	22278996 264508 264600 264602 264603
		oil3122359lspiO33123/LEU2_MYCLE - 3-	Aconitace family (aconitate		DEAGOE SSETTOS SEARCE SCAROS, ESTUGS,
		ISOPROPYLMALATE DEHYDRATASE LARGE SUBUNIT	Acumase ranny (acomitate hydratase)		264605, 3365/023, 264565, 264486
		(ISOPROPYLMALATE ISOMERASE) (ALPHA-IPM			
7	1007 1007 1000	ISCINILITY OF (IT MI)			
4	000/3361 (427, 428)	Novel Protein sim. GBank	Contains protein domain (PF00118) - eph	eph	264600, 264693
		GINTEZSESPIP19421 CH60 COXBU - 60 KD	TCP-1/cpn60 chaperonin family		
		(HEAT SHOCK PROTEIN CYNOU) (GRUEL PROTEIN)			
215	14973283 (429, 430)			UNCLASSIFIED	264629
216	80177716 (431, 432)	Novel Protein sim. GBank gi 3417297 (AC002310) -	Contains protein domain (PF00096) - dna_rna_bind	dna_rna_bind	264448
217	79603634 (433 434)	Novel Design aim Cook	Zinc Inger, C2H2 type		
: 	(101 (201)	dil2506924 solP49754 VP41 HTMAN - VACTOLAR			264508
		ASSEMBLY PROTEIN VPS41 HOMOLOG (S53)			
218	80258475 (435, 436)	Novel Protein sim. GBank		rnapolymerase	264594
		gij1173288 sp P38106 RSEA_ECOLI - SIGMA-E FACTOR NEGATIVE REGUI ATORY PROTFIN			
219	20438797 (437, 438)	Novel Protein sim. GBank gil1781097 emb CAB06231		svnthase	264604
Š	0077 000007	(Z83864) gltB [Mycobacterium tuberculosis]			
077	13488372 (438, 440)	Novel Protein sim. GBank gi 2984703 (AF052427) - unknown Trypanosoma cruzil		nucleaseinhib	264689
221	11287498 (441, 442)	Novel Protein sim. GBank gil4587313 dbj BAA76709.1 -		UNCLASSIFIED	264555
		(AB025248) alpha-1,2-mannosidase [Bacillus sp. M-90]			

222	79862802 (443 444)	Novel Drotein sim Coast aitt 07735812mbl ADA70401			3010000 000100 100100
		(29270) hypothetical protein Rv0143c [Mycobacterium tuberzulosis]		ONCEASSIFIED	204000, 204709, 33090423
223	83053869 (445, 446)			UNCLASSIFIED	264906 264907 264603
224	79557920 (447, 448)				264684, 264693
225	79559541 (449, 450)	Novel Protein sim. GBank gi[2274851 dbj BAA21515 - (D64159) 3-7 cene product (Homo saciens)		UNCLASSIFIED	264692
226	79172397 (451, 452)			UNCLASSIFIED	22278998, 264112, 33657023, 263981
227	81777196 (453, 454)			UNCLASSIFIED	35695917 264636 264907
228	79872285 (455, 456)				264768, 264907, 264908, 264692, 264593,
229	79838266 (457, 458)				26490B 264910
230	11013209 (459, 460)			UNCLASSIFIED	264631
231	20622207 (461, 462)	Novel Protein sim. GBank gij1835114jembjCAA71733j - (Y10744) homoserina Olacabultanefarase II ophosica			264906, 264600, 264603, 264692
		meyen]			
232	80055035 (463, 464)			UNCLASSIFIED	264600, 264603, 264605, 264687, 264769
233	80063054 (465, 466)	_	Contains protein domain (PF00449) - hydrolase Urease	hydrolase	264604
234	7523998 (467, 468)	Novel Protein sim. GBank gij3510505 (AF030881) - pol polyprotein (Fuou rubribes)		UNCLASSIFIED	264369
235	80203671 (469, 470)			UNCLASSIFIED	264106
236	78940001 (471, 472)	Novel Protein sim. GBank gi[2104609 emb CAB08805 - 1725398) PrkA (Mycharlerium Ingrae)		carboxylase	264905
237	11755273 (473, 474)				264681
ı	79461401 (475, 476)			IINCI ASSIFIED	26463Q
1	82435190 (477, 478)	Novel Protein sim GBank	Contains protein domain (DE00037)		264006 765040 264603 264763 264683
- 1		IJ_HAEIN - HYPOTHETICAL	4Fe-4S ferredoxins and related ironsularly sulfur cluster binding domains.		264636, 264638, 264486
240	21635575 (479, 480)	ABC		transport	264259, 264769
241	80377307 (481, 482)	Novel Protein sim. GBank gij3875920jemb CAB04111 - (Z81503) predicted using Genefinder; similar to collagen; cDNA EST EMBL:D65450 comes from this gene; cDNA EST EMBL:D68888 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	264908, 264909, 264764, 264639
242	82148454 (483, 484)			UNCLASSIFIED	264489, 264907, 264908, 264511, 264760, 264764, 264692, 264635, 264637
243	79633207 (485, 486)			UNCLASSIFIED	264906
244	80248682 (487, 488)	Novel Protein sim. GBank gi 2624302 emb CAA15575 - (AL008967) ald [Mycobacterium tuberculosis]		dehydrogenase	264600, 264602, 264605, 264769, 264689
245	79863543 (489, 490)	Novel Protein sím. GBank gil2920625 (AF044499) - vgrE protein [Escherichia coli]		UNCLASSIFIED	264907, 264758
246	79162929 (491, 492)	Novel Protein sim. GBank gi[5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]	Contains protein domain (PF01106) - NifU-like domain		264637, 18108381, 18108387, 264565

247	79873185 (493, 494)	Novel Protein sim. GBank pil1839006lemblCAB066481 -		kingea	764000 264601 26506422 18400287
		(Z85982) argB [Mycobacterium tuberculosis]		9	ייסטטיין דייסטטיין, טיסטטיין, ייסטטטיין, ייסטטטטיין, ייסטטטטיין, ייסטטטטיין, ייסטטטטיין, ייסטטטטיין, ייסטטטטיין, ייסטטטטטיין, ייסטטטטטטיין, ייסטטטטטטיין, ייסטטטטטטטטטטטטטטטטטטטטטטטטטטטטטטטטטטט
248	80488983 (495, 496)	Novel Protein sim. GBank		synthase	35696286, 264907, 264511, 264602, 264768,
		gi 1168574 sp P42464 ATPB_CORGL - ATP SYNTHASE BETA CHAIN			264688, 265021, 35695855, 18108385
249	79764645 (497, 498)			UNCLASSIFIED	264907, 264910, 265011, 264762, 264636
220	79619980 (499, 500)				21906768, 264692
221	84359489 (501, 502)			UNCLASSIFIED	52645156, 29331822, 29331824, 52644045, 265018, 21906765, 21906768, 265020, 27486261, 27486265, 35695763, 18108376, 26565
252	79737756 (503, 504)	Novel Protein sim. GBank gij3327166 dbj BAA31651 - (AB014576) KIAA0676 protein [Homo saplens]			264685, 264687, 264632
253	20443124 (505, 506)	Novel Protein sim. GBank gij3036880[emb[CA418513] - (AL022374) putative ATP-dependent DNA helicase [Streptomyces coelicolor]		helicase	264604
254	80027421 (507, 508)	Novel Protein sim. GBank gi 3915488 sp 034961 yJMB_BACSU - HYPOTHETICAL SYMPORTER IN COTT-RAPA INTERGENIC REGION		UNCLASSIFIED	264508, 264906, 264602, 264687, 265021, 264486
255	11398315 (509, 510)	Novel Protein sim. GBank giļ1665720 dbj BAA04134 - (D17312) diarrheal toxin [Bacillus cereus]		UNCLASSIFIED	264593
256	80028158 (511, 512)	Novel Protein sim. GBank gil465787IspIP34422 YL31_CAEEL - HYPOTHETICAL 86.0 Prolyl oligopeptidase family KD PROTEIN F4489.1 IN CHROMOSOME III	Contains protein domain (PF00326) - peptidase Protyl oligopeptidase family	peptidase	264602, 264692
257	20289262 (513, 514)	Novel Protein sim. GBank gi[1172039 sp P42315 SCOA_BACSU - PROBABLE SUCCINYL-COA:3-KETOACID-COENZYME A TRANSFERASE SUBUNIT A (SUCCINYL COA:3-OXOACID COA-TRANSFERASE) (OXCT A)	Contains protein domain (PF01144) - transferase Coenzyme A transferase	transferase	264605
528	20459464 (515, 516)	Novel Protein sim. GBank gil3127836 emb CAA18902 - (AL023496) hypothetical protein [Streptomyces coelicolor]		UNCLASSIFIED	264604
259	79910152 (517, 518)				264681, 264686, 264692
3 5	2027 5437 (519, 520)	Novel Dratein eim CBank	ONCLASSIFIED (NICLASSIFIED		264692, 264556
	()	gi 12376 1sp P24221 HUTH_STRGR - HISTIDINE AMMONIA-LYASE (HISTIDASE)	Contains protein domain (PF00221) - Phenylalanine and histidine ammonia lyases		764500
262	80189317 (523, 524)			UNCLASSIFIED	265017, 264369
Z 293	88095045 (525, 526)	Novel Protein sim. GBank gil3924708 emb CAA84646 - (Z35597) Weak similarity with sea squirt nidogen precursor		UNCLASSIFIED	264488, 264905, 264906, 264907, 264908, 264909, 264512, 264910, 264758, 264596,
		protein (blastp score 71); cDNA EST EMBL: T02069 comes from this gene; cDNA EST EMBL: D76135 comes from this			264604, 265019, 264605, 264760, 18108351, 264763, 264764, 264288, 264768, 264768
		gene; cDNA EST EMBL:D73147 comes from this gene; cDNA EST EMB.			264769, 264691, 264692, 264693, 264628, 264694, 264694, 264697, 264699, 264699, 264698, 264694, 2646404, 264644, 264644, 2646444, 264644, 264644, 264644, 264644, 264644, 264644, 264644, 2646
					204034, 204033, 204030, 204030, 204030, 264639
79	87370826 (527, 528)	Novel Protein sim. GBank gi 3043734 dbj BAA25531 - (AB011177) KIAA0605 protein [Homo sapiens]	Contains protein domain (PF00047) - protease Immunoglobulin domain	protease	264259, 264908, 21906754, 265018, 265019, 265020

	ı			
565		Novel Protein sim. GBank gil4589624 dbj BAA76834.1 - (AB023207) KIAA0990 protein [Homo sapiens]	Kinase	264488, 35696286, 29331824, 56182181, 35696052, 264508, 264905, 264906, 264907, 66712502, 24908, 264909, 264511, 24512, 264910, 264502, 224596, 264511, 24512, 26491386, 264600, 265017, 264603, 264604, 264288, 264760, 264762, 264762, 264681, 24966765, 21906765, 21906767, 21906769, 265020, 264691, 33657023, 33657109, 33657102, 264631, 264638, 264636, 264636, 264586, 264536, 2644386, 264536, 264536, 2644386, 264536, 264536, 2644386, 264536, 264536, 2644386, 264536, 264436, 264536, 264436, 264536, 264436, 264536, 264436, 264536, 264436, 264536, 264436, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 264436, 26456, 2644486, 26456, 2644486, 26456, 2644486, 26456, 2644486, 26456, 2644486, 26456, 2644486, 26456, 2644486, 26456, 2644486, 26456, 2644486, 26456, 26456, 264486, 26456, 26456, 264486, 26456, 26456, 264486, 26456, 26456, 264486, 26456, 26456, 26456, 26456, 26456, 26456, 26456, 26450
266	79588075 (531, 532)			264600
267	11362222 (533, 534)		UNCLASSIFIED	264828
268	79909566 (535, 536)		UNCLASSIFIED	264687, 264769, 264689
269	80025810 (537, 538)		UNCLASSIFIED	264602
270	84361144 (539, 540)	Novel Protein sim. GBank pi4507367IrefiNP 003182 1h7ARS - threnow-IRNA	UNCLASSIFIED	264693
		sproduced in the second of the		
112	79552301 (541, 542)		UNCLASSIFIED	264909, 264693
272	9674778 (543, 544)	Novel Protein sim. GBank	synthase	264908
		gil4980736 gb AAD35331.1 AED0170 - (AE001707) glucose- 1-phosphate adenylytransferase [Thermotoga maritima]		
273	12840694 (545, 546)	Novel Protein sim. GBank	UNCLASSIFIED	264688
		gi 1168224 sp P44569 5NTD_HAEIN - PROBABLE 5'- NUCLEOTIDASE PRECURSOR		
274	39524246 (547, 548)			264564
275	82787041 (549, 550)	Novel Protein sim. GBank gi]3253159 (AF005355) - Iranslation initiation factor eIF2C [Oryctolagus cuniculus]	UNCLASSIFIED	264907, 264908, 264909, 264766, 264768, 264691, 264632, 264636
276	86671073 (551, 552)	Novel Protein sim. GBank gil134920lsplP21997[SSGP_VOLCA - SULFATED SLIREACE GI YCOPROTEIN 185 (SSG 185)		265008, 60432229
277	80079735 (553, 554)	Novel Protein sim. GBank	ribosomalprot	264600, 18108387
		gi 129021 sp P20964 0BG_BACSU - SPO0B-ASSOCIATED GTP-BINDING PROTEIN		
278	12966947 (555, 556)		UNCLASSIFIED	264689
279	95292719 (557, 558)	Novel Protein sim. GBank gij79839 pir S03812 - uvrB	nuclease	264508, 264604, 21906764, 264638, 264557, 264404
280	5603617 (559, 560)			264259
281	80249599 (561, 562)	Novel Protein sim. GBank		18108392, 264634, 264555, 264556, 264557,
		gi3123160jspjQ18964JYLN2_CAEEL - HYPOTHETICAL 46.2 KD TRP-ASP REPEATS CONTAINING PROTEIN		264558
cac	10500000 (562 564)	DZ013.2 IN CHROMOSOME II	TINCI ACCICIED	265010
	10090002 (303, 304)		ONCEASSIFIED	810007
283	20614211 (363, 360)		UNCLASSIFIED	264555

è	ł			-	
404	(905, 206) (907, 308)	Novel Protein Sim. GBank gi[2429094 (U58632) - acely!	Contains protein domain (PF00300) - UNCLASSIFIED	UNCLASSIFIED	35696052, 29331828, 264508, 264905.
		xylan esterase; AxeA [Thermotoga neapolitana]	Phosphoglycerate mutase family		264600, 264602, 264605, 264682, 264764,
					56181562, 21906764, 18108376, 264636,
					264559, 18108387
285	8757940 (569, 570)			UNCLASSIFIED	264603
286	80503235 (571, 572)	Novel Protein sim. GBank gi 2072674 emb CAB08305 -	Contains protein domain (PF00270) - ATPase_associated 35696052, 264769, 264638	ATPase associated	35696052, 264769, 264638
		(Z95120) rhIE [Mycobacterium tuberculosis]	DEAD/DEAH box helicase	1	
287	12745521 (573, 574)			UNCLASSIFIED	264689
288	(20756502 (575, 576)	Novel Protein sim. GBank gil765323lbbs 157676 - (S74439)		collagen	264557
		silk fibroin heavy chain (C-terminal) [Bombyx		•	
		mori=silkworms, Peptide Partial, 633 aa] [Bombyx mori]			
289	80043804 (577, 578)	Novel Protein sim. GBank gij1870009 emb CAB06860 -	Contains protein domain (PF00440) - ribosomalprot	ribosomalprot	264593, 264600
		(Z92539) hypothetical protein Rv1019 [Mycobacterium	Bacterial regulatory proteins, tetR		
		[tuberculosis]	family		
290	80430175 (579, 580)			UNCLASSIFIED	264768
291	20747431 (581, 582)	Novel Protein sim. GBank		UNCLASSIFIED	264601
_		Init 2506664 Iso IDA0120 IVOC ECOLL SO A DECITE IN IN			
		PHENOMOTISM TO THE PROPERTY OF			
200	ROD52555 (583 584)	Novel Protein rim Chark Allenda 1 20015			200700
707	decox 200 (200, 204)	INOVER FIDICIAL SIGN. GERNIK GIJOZOTOZ (L.39010) -		UNCLASSIFIED	264605
_		mitochondrial glutamyl-tRNA synthetase [Saccharomyces			
		cerevisiae]			
293	80062519 (585, 586)	Novel Protein sim. GBank		helicase	264909, 264605, 264687, 264689, 264692
		Dil17180651snIP5352811WRD MYCLE - PLITATIVE DNA			
		HELICASE II HOMOLOG			
294	79830303 (587, 588)	Novel Protein sim. GBank	Contains protein domain (PF00008) - oncogene	oncogene	35696052, 264906, 265011, 264628,
		gij117422 sp P10040 CRB_DROME - CRUMBS PROTEIN	EGF-like domain		55811576
285	79444180 (589, 590)	Novel Protein sim GBank nil1181619IdhilB44115651			52644507 20331822 264503 265020
	(22.2)	(D82364) a variant of TSC-22 (Gallus natural			25044301, 233310££, 20433£, 2030£0,
596	79607076 (591, 592)	Novel Protein sim. GBank qi 3649789 dbi BAA33403 -		synthase	264508
		(AB012226) SecA [Vibrio alginolyticus]			
297	79631297 (593, 594)	Novel Protein sim. GBank gil5689967 emblCAB52004.11 -		UNCLASSIFIED	264905, 264687, 264638
		(AL 109663) putative membrane protein [Streptomyces			
		coelicolor A3(2))			
298	80418898 (595, 596)			UNCLASSIFIED	264905, 264691, 264639, 264766

	95293298 (597, 598)	Novel Protein sim. GBank gi 220637 db BAA01477 - (D10627) zinc finger protein [Mus musculus]	Contains protein domain (PF00096) - Zinc finger, C2H2 type		264488, 263994, 56994075, 22278997, 22278998, 22278999, 20281099, 29331824, 29331825, 60432289, 29331824, 29331828, 264905, 264907, 264908, 2644045, 264909, 264901, 265409, 264908, 2644045, 264909, 264511, 265008, 264910, 264595, 264596, 264511, 265008, 264910, 264595, 26468, 265018, 265019, 264764, 264288, 264065, 21906762, 21906762, 21906762, 21906762, 24629, 5811576, 35695855, 264629, 33657109, 27486261, 18108370, 264628, 264629, 264639, 3657109, 246535, 264638, 22279000, 22279002, 264565, 264566, 264567
ဓ္ဓ	20711340 (599, 600)			UNCLASSIFIED	264602
301	13511332 (601, 602)	Novet Protein sim. GBank gil 145922 (M20981) - iron dictirate transport protein precursor [Escherichia coli]		transport	264687
305	9875260 (603, 604)	Novel Protein sim. GBank gi 1174661 sp P44594 TGT_HAEIN - QUEUINE TRNA- RIBOSYLTRANSFERASE (TRNA-GUANINE TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME)			264908
303	79574895 (605, 606)				264689
304	20711344 (607, 608)	Novel Protein sim. GBank gil67985 pir HJNVAV - helicase (EC 3.6.1) - Autographa californica nuclear polyhedrosis virus		helicase	264602
305	80412520 (609, 610)	Novel Protein sim. GBank gij728867 sp P40602 APG_ARATH - ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR			264763
906	8515876 (611, 612)	Novel Protein sim. GBank gil1657554 gb AAB18082.1 - (U73857) hypothetical protein [Escherichia coli]		UNCLASSIFIED	263978
307	80222901 (613, 614)			UNCLASSIFIED	265010, 21906768, 265020, 18108374, 263977
308	80064305 (615, 616)	Novel Protein sim. GBank gi 1710612 sp Q10793 RNH2_MYCTU - PROBABLE RIBONUCLEASE HII (RNASE HII)	Contains protein domain (PF01351) - nuclease Ribonuclease HII	nuclease	264910, 264600, 264605, 264687, 264689, 264638, 18108387
308	80504136 (617, 618)	Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]			264769
330	80053616 (619, 620)				264603
311	11090659 (621, 622)	Novel Protein sim. GBank gil 1144522 (U34957) - phosphoribosylaminoimidazolesuccinocarboxamide synthase (Mycobacterium tuberculosis)		synthase	264602
312	80054347 (623, 624)			UNCLASSIFIED	264566
313	80046168 (625, 626)				264603, 264567

314	87645112 (627, 628)	Novel Protein sim. GBank gi 3661583 (AF092175) - ikaros	Contains protein domain (PF00320) - dna rna bind	dna rna bind	264259 60432289 29331828 264905
		[Danio rerio]	GATA zinc finger	1	264906, 264908, 264909, 265008, 264910,
					60432229, 33657402, 60433438, 33109954, 1 265011, 265017, 264603, 265018, 264288,
					264766, 264692, 35695763, 264628, 264629,
315	82356091 (629 630)	Novel Protein sim GRank gil165262014hilB0.0175401			264639, 60170394, 22279002, 264300
<u> </u>	(200)	(D90907) pyridine pucleotide transhydrogenase heta			204300, 204000, 204/02, 204001, 204/06, 52644330, 264636
		subunit (Synechocystis sp.)			264638, 264486
316	79911071 (631, 632)			UNCLASSIFIED	264693
317	20466944 (633, 634)	Novel Protein sim. GBank			264605
		gil118244 sp P24176 DAPE ECOLI - SUCCINYL-			
		DIAMINOPIMELATE DESUCCINYLASE (SDAP)			
318	94141836 (635, 636)	Novel Protein sim. GBank	Contains protein domain (PF00526) - Itransport	transport	264908, 264909, 264910, 264593, 264594
		gil4680229lgbjAAD27583.1 AF11827 - (AF118274) DNb-5	Dictyostelium (slime mold) repeats		264760, 264288, 264768, 264769, 21906769
		[Homo sapiens]			264691, 264693, 264628, 65274791, 264635,
					264636, 264638, 83373044, 22279002,
5	1000 1000 00000011				264566
<u> </u>	11/289300 (637, 638)	Novel Protein sim. GBank gi 1149693 emb CAA60220 -		transport	265018
5	1010 000/ 35050308	(Accepta) ruse (Clostingiam) perinngens			
350	13527675 (639, 640)	Novel Protein sim. GBank		synthase	264687
		griza i iussispicuosi 4 jei GC MYCIU - GLUCUSE-1-			
		PHOSPHATE ADENYLYLTRANSFERASE (ADP.			
		GLUCOSE SYNTHASE) (ADP-GLUCOSE			
		PYROPHOSPHORYLASE)			
321	94134387 (641, 642)	Novel Protein sim. GBank gi 1680716 (U68234) - all-trans-		cyto450	264509, 264906, 264907, 264908, 265009.
		retinoic acid 4-hydroxylase [Danio rerio]			264596, 264764, 264628, 264634, 264635,
					264638, 264639, 83373044, 264567
322	66489053 (643, 644)	Novel Protein sim. GBank gil1160355 (U33058) - UNC-89		UNCLASSIFIED	55811150, 264691, 60431528, 55810764
		[Caenomapditis elegans]			
323	94653725 (645, 646)			UNCLASSIFIED	264488, 265009, 264593, 264628, 264635
324	79174383 (647, 648)				264687
325	79862691 (649, 650)				264693
328	28774974 (651, 652)			UNCLASSIFIED	264288, 18108385
327	79776267 (653, 654)	Novel Protein sim. GBank gi 451544 (U04267) - proline-rich			264488, 264905, 264509, 264910
328	RODESOND (REE BEEN	con trail broad brain parparensel		41.100	
	40472004 (053, 050)			UNCLASSIFIED	264592
329	10173821 (657, 658)			UNCLASSIFIED	264510
330	86597767 (659, 660)	Novel Protein sim. GBank gil4191358 (AF087825) - claudin-		UNCLASSIFIED	264259, 264908
		7 [Mus musculus]			
331	79754888 (661, 662)	Novel Protein sim. GBank gi 80741 pir S20912 - regulatory protein whi8 - Strentomycas coelicolor		transcriptfactor	264910, 264687, 264689, 264636, 264567
332	80071440 (663 664)	Novel Protein eim Chank			TOTAL DESTRUCTION CONTROL CONTROL
	(100)	gil114049lsp P19480lAHPF SALTY - ALKYL		reductase	330904£3, 204030, 204036, 204303
		HYDROPEROXIDE REDUCTASE SUBUNIT F (ALKYL			
200	100000000000000000000000000000000000000	HYDROPEROXIDE REDUCTASE F52A PROTEIN)			
3	13009000 (000, 000)				264687

264905, 264600, 264604, 264486	264907, 264603, 264693, 18108374, 264636, 18108387	265009, 264766, 264686	02	264569, 18108397, 22278998, 29331822,	20281099, 29331824, 56182181, 66714117. 29331825, 35696052, 29331828, 264508.	264509, 264905, 264906, 264907, 264908.	264909, 265006, 265008, 264910, 265009,	264758, 55812038, 65274444, 265011,	87168559, 265017, 265018, 265019, 264760, 55811150, 264681, 264762, 18108351.	264682, 264764, 264765, 264685, 264686.	264768, 52644229, 264689, 55811957,	35695917, 264692, 264693, 264628,	18108370, 18108374, 55811576, 35696423,	35695855, 264635, 264555, 264636, 264556,	264637, 264557, 18108380, 264638, 264558,	264639, 18108381, 83373044, 18108385,	8/108315, 00432113	193	264905, 265019, 264769, 18108374					505	264603 264604			264605	500, 264605, 264768, 18108370,	18108374, 35695855	593		905	
		UNCLASSIFIED 2650	264602		2028	2645	2649	2647	8716	2646	2647	3569	1810	3569	2646	2646		eph 264593	ubiquitin 2649					somerase 264605	2646				UNCLASSIFIED 2646	181	histone 264593		dehydrogenase 264905	
Contains protein domain (PF00208) - or Glutamate/Leucine/Phenylalanine/Valine dehydrogenase				Contains protein domain (PF00249) - nucl_recpt	Myb-like DNA-binding domain														1					Contains protein domain (PF00290) - Isomerase	Iryptophan synthase alpha chain								Contains protein domain (PF00208) - dehydrogenase (Shitamate) euclie/Phenylalanine/va	line dehydrogenase
n n	Novel Protein sim. GBank gi[2193938]emb CAB09602 - (296800) gipQ2 [Mycobacterium tuberculosis]			Novel Protein sim. GBank	gi]5454074/ref NP_006303.1 pSMRT - silencing mediator for Myb-like DNA-binding domain refinoid and thyroid hormone recentors													Novel Protein sim. GBank gi 4001713 dbj BAA35087.1 - (AB015879) DnaK [Porphyromonas gingivalis]	Novel Protein sim. GBank	gi[2842699]sp[Q92353]UBPC_SCHPO - PUTATIVE	VIBIOLITIN THIS ESTERASEVILIBIOLITIN-SPECIFIC	DESCRIPTION OF THE PROPERTY OF	ENZYME)	Novel Protein sim. GBank gi 2688580 (AE001166) -	Conserved hypothetical protein Borrella burgdorferi	(Y09452) Yed j hypothetical protein [Pseudomonas	syringae)		Novel Protein sim. GBank gi[2117275 emb CAB09104 -	(Z95618) hypothetical protein Rv0807 [Mycobacterium tuberculosis]	Novel Protein sim. GBank	gij3023317jspjQ48935jaPHA_MYCRA - ACETYLPOLYAMINE AMINOHYDROLASE	Novel Protein sim. GBank gil4239787 emb CAA75437 -	aeruginosa]
80230771 (667, 668)	80057026 (669, 670)		11090829 (673, 674)	95413134 (675, 676)													╗	11398513 (677, 678)	80504149 (679, 680)					11075198 (681, 682)	900E410E (893 894)	, con (con; con)		20466792 (685, 686)	80428870 (687, 688)		80258853 (689, 690)		79831058 (691, 692)	
334	335	336	337	338														339	8				_	341	CFC	\$	i	343	344		345	_	346	

347	79158195 (693 694)	Novel Protein sim Chapk			
1		gij731675 spiP38795 YHN4_YEAST - HYPOTHETICAL 80.7 KD PROTEIN IN ERG7-NMD2 INTERGENIC REGION		UNCLASSIFIED	263981
348	80020208 (695, 696)			transport	264602, 18108351, 18108387
349	17282112 (697, 698)				265007
320	80502370 (699, 700)	Novel Protein sim. GBank gij3261599 emb CAB00917 - (Z77137) hypothetical protein Rv1277 [Mycobacterium tuberculosis]		nuclease	265009, 264769, 264689, 18108370
351	80501805 (701, 702)	Novel Protein sim. GBank gi[2959367 emb CAA17921 - (AL022117) hypothetical protein [Schizosaccharomyces pombe)		glycoprotein	264769, 264905, 264908
352	11611585 (703, 704)	Novel Protein sim. GBank gil4416302[gb AAD20307] - (AF105716) copia-type pol polyprotein (Zea mays)		protease	264595
	80061653 (705, 706)	Novel Protein sim. GBank gi[1174887 sp P42873 URE1_STAXY - UREASE ALPHA SUBUNIT (UREA AMIDOHYDROLASE)	Contains protein domain (PF00449) - UNCLASSIFIED Urease	UNCLASSIFIED	264604
	56626130 (707, 708)			INCI ASSIEIED	264628
355	80046344 (709, 710)			UNCLASSIFIED	264909 264595 264683 22279002
	80043835 (711, 712)	Novel Protein sim. GBank gil115157 sp P16574 BVGA_BORPE - VIRULENCE FACTORS PUTATIVE POSITIVE TRANSCRIPTION REGULATOR BVGA	Contains protein domain (PF00072) - transcriptfactor Response regulator receiver domain	transcriptfactor	264909, 264591, 264592
357	80070566 (713, 714)	Novel Protein sim. GBank gil497637 (J03939) - cytochrome oxidase d subunit I [Escherichia coll]		oxidase	264605
358	37032756 (715, 716)	Novel Protein sim. GBank gi[2290990 (AF006000) - Brg1 [Bordetella pertussis]		UNCLASSIFIED	264768
8	80501488 (717, 718)			UNCLASSIFIED	264604, 264769
9	80026748 (719, 720)			UNCLASSIFIED	264594
ę,	60584075 (721, 722)	Novel Protein sim. GBank gij3510639 (AF049344) - UDP. GalNAc:polypeptide N-acetylgalactosaminyltransferase T5 [Rattus norvegicus]		transferase	22278996, 264259, 29331822, 29331824, 264605, 55811957, 265022
362	13089485 (723, 724)	Novel Protein sim. GBank gil113764 sp P25718 aMY1_ECOLI - ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)		amylase	264688
363	79750145 (725, 726)				264566
364	82443593 (727, 728)	Novel Protein sim. GBank gi[2829816 sp P95171 NUOK_MYCTU - NADH DEHYDROGENASE I CHAIN K (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 11) (NUO11)	Contains protein domain (PF00420) - dehydrogenase NADH-ubiquinone/plastoquinone oxidoreductase chain 4L	dehydrogenase	264638, 264602, 264604, 264508, 264762, 264638, 264486

Į	80056153 (755 756)			UNCLASSIFIED	265008 264555
10503437	(757 758)	Novel Protein sim GBack ail1076013 lairt[A40030 - carB	Contains protein domain (DE00389) - (September 1989)		264760
r tooppo	(901,101)	cobacterium bosis (efrain BCC)	Cartamovi photomical (cr cocco)		60.407
			(CPSase)		
9006093	80060937 (759, 760)	ein sim. GBank gij216556 dbj BAA02174 - glucose dehydrogenase [Escherichia coli]	Contains protein domain (PF01011) - dehydrogenase PQQ enzyme repeat	dehydrogenase	264604
1176902	11769027 (761, 762)			UNCLASSIFIED	264684
3005437	80054377 (763, 764)				264592
832590	83259025 (765, 766)	Novel Protein sim. GBank gij3327136[dbj]BAA31636] - (AB014561) KIAA0661 protein [Homo sapiens]			264595, 265017, 265021, 264638, 87168518, 22279002
953142	95314255 (767, 768)			UNCLASSIFIED	264259, 29331822, 60432289, 29331827, 264288, 264766, 263967, 65274791, 35695855, 263981, 83373044, 264567
102376	10237679 (769, 770)				264692
796334	79633434 (771, 772)	Novel Protein sim. GBank gi 1073456 pir S47810 - probable Contains protein domain (PF00465) - dehydrogenase alcohol dehydrogenase (EC 1.1.1.1) - Escherichia coli dehydrogenases	Contains protein domain (PF00465) - Iron-containing alcohol dehydrogenases	dehydrogenase	264906
179606			Contains protein domain (PF01841) - UNCLASSIFIED Transglutaminase-like superfamily	UNCLASSIFIED	264760
877413	87741376 (775, 776)	74863.1 -	Contains protein domain (PF00646) -	homeobox	35696286, 264905, 66712502, 60432229,
			F-box domain.		264593, 60433356, 264686, 264688, 21906765, 264691, 22279000, 264482
793169	79316971 (777, 778)			UNCLASSIFIED	18108394, 22278996, 264630, 264556, 22279002
800799	80079949 (779, 780)			UNCLASSIFIED	264600
765730	7657302 (781, 782)	Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6]			264482
797960	156 (783, 784)			UNCLASSIFIED	264908
332060	33206031 (785, 786)	Novel Protein sim. GBank gi[3378523]emb CAA08867] - (AJ009832) cyclomaltodextrinase glucanotransferase [Thermotoga neapolitana]		synthase	264602, 21906764
10104	10104463 (787, 788)				264693
802290	80229010 (789, 790)			UNCLASSIFIED	264508, 264563
20436;	20436224 (791, 792)	gi 2677780 (U70327) - unknown	Contains protein domain (PF00047) - struct Immunoglobulin domain	struct	264556
80417(80417014 (793, 794)	Novel Protein sim. GBank gil4507909 ref NP_000368.1 pvVAS - Wiskott-Aldrich syndrome (ecezema-thrombocytopenia)			265007, 265009, 264508, 264556, 264629, 264766
91230	91230517 (795, 796)	Novel Protein sim. GBank gil 1518458 (U45998) - mitochondrial solute carrier [Onchocerca volvulus]	Contains protein domain (PF00153) - transport Mitochondrial carrier proteins	transport	18108398, 22278995, 22278996, 56994075, 22278999, 264259, 29331824, 29331826, 264905, 264908, 265007, 265008, 265009, 264288, 264766, 21906765, 21906766, 21906767, 266020, 265021, 33657023, 33657109, 264228, 35696423, 35695855, 264952, 18108391

399	80055278 (797, 798)	Novel Protein sim. GBank gij3358091 ldbj BAA31995 - (AB015974) glycerol kinase [Pseudomonas tolaasii]	Contains protein domain (PF00370) - kinase FGGY family of carbohydrate	kinase	264592, 264595
			kinases		
0	94117490 (799, 800)	Novel Protein sim. GBank gil72835[sp[P39192]ALU5_HUMAN - !!!! ALU SUBFAMILY Leucine Rich Repeat SC WARNING ENTRY !!!!	Contains protein domain (PF00560) - cadhenn Leucine Rich Repeat	cadherin	18108394, 56182575, 22278995, 22278997, 22278999, 264259, 29331824, 265006, 265007, 265009, 60432229, 33657402, 21906754, 265010, 265017, 265018, 265019, 18108357, 21906765, 265021, 265022, 264691, 264692, 33657023, 18108370, 65274791, 264634, 264636, 264636,
401	11397491 (801, 802)	Novel Prolein sim. GBank gil4928292[gb]AAD33527.1 AF13211 - (AF132117) FhuA [Staphylococcus aureus]		transport	60170394, 56182323 264594
70	824.20.284 (803. 804)	Novel Proten sm. GBank gij5689487(dbj BAA83027.1 - (AB028998) K.AA1075 protein [Homo sapiens]	Contains protein domain (PF00017) - phosphatase Src homology domain 2		65274572, 56182575, 35696286, 22278996, 22278998, 264093, 264259, 29331822, 29331824, 29331825, 29331826, 60433289, 264099, 264093, 264591, 265006, 264907, 264909, 265006, 264511, 265007, 265008, 264501, 264691, 264691, 2046354, 2644296, 265019, 264681, 18108351, 264682, 264448, 265019, 264681, 18108351, 264682, 26448, 264288, 264684, 264766, 264767, 294686, 21906766, 21906766, 21906769, 265021, 265022, 264629, 18108370, 264639, 18108370, 264639, 26273902, 264557, 264639, 18108386, 26273902, 264561, 264561, 264565, 264566, 26273902, 264561, 264565, 264566, 264565, 264566, 26273902, 264561, 264561, 264565, 264566, 26273902, 264561, 264561, 264565, 264566, 264666,
Ş	80439913 (805, 806)			UNCLASSIFIED	264768, 264632, 264639, 264563
\$ \frac{4}{5}	11809865 (807, 808) 79471280 (809, 810)	Novel Protein sim. GBank gi 2661649 emb CAA15755 -		polymerase	264682 265009 264682
4 9 9	79634172 (811, 812)	(AL009198) dnaE2 [Mycobacterium tubercutosis]	Contains profess domain (BE00150)		COOPED LICEOTE
			Pancreatic hormone peptides	UNCLASSIFIED	18108357, 264693
9	80478229 (813, 814)			UNCLASSIFIED	264769
§ §	80079956 (815, 816)				264600
£03	5640527 (817, 818)	Novel Protein sim. GBank gij3047117 (AF058919) - similar to ATP-dependent RNA helicases [Arabidopsis thaliana]			264259

0	95357496 (819, 820)	Novel Protein sim. GBank gil475016 dbj BAA06184 - (D29801) Unknown [Mus musculus]		UNCLASSIFIED	264489, 52646365, 52646842, 56181686, 35696286, 52645080, 29331822, 29331824, 5618181, 29331824, 5618181, 29331824, 5618181, 29331824, 5618181, 29331824, 5618181, 29331824, 5618181, 29331824, 5618181, 29331824, 2649163, 2649104, 265907, 264908, 2644906, 265007, 264517, 265008, 264910, 39557402, 264764, 265018, 265011, 265017, 264604, 265018, 55811150, 265011, 265012, 264016, 265012, 265020, 265012, 264364, 264682, 264629, 33657023, 65274620, 33657109, 33657182, 27486281, 35693762, 264632, 264632, 264633, 264633, 264633, 264633, 264633, 264633, 264633, 264631, 2664315, 264634, 264633, 264633, 264633, 264631, 26643113, 22279002, 264563, 264565, 264631, 264331, 264533, 264565, 264631, 264331, 264533, 264565, 264431, 264533, 264633, 264
411	80501670 (821, 822)			OSCIENCE A CONTRACTO	264780
412	80241662 (823, 824)	The state of the s			264907 264910 263973 22279002
413	11076446 (825, 826)	Novel Protein sim GRank gil 32617841embl Co B080071			201301, 203310, 203313, 22213002
2	1101 0410 (040)	(295558) htpX [Mycobacterium tuberculosis]		epn	264 603
414	82050554 (827, 828)	Novel Protein sim. GBank		dehydrogenase	18108374, 264760, 264769, 264602, 264638,
		gi 129036 sp P20707 ODO1_AZOV1 - 2-OXOGLUTARATE DEHYDROGENASE E1 COMPONENT (ALPHA- KETOGLUTARATE DEHYDROGENASE)			264603, 264909, 264805
415	84453144 (829, 830)	Novel Protein sim. GBank		UNCLASSIFIED	264908, 87168518
		gil4868350lgb AAD31273.1 AF13202 - (AF132025) rhophilin Drosophila melanogaster]			
416	80402775 (831, 832)	Novel Protein sim. GBank gil2555172 (AF025543) - ArcC; carbamate kinase IRbizobium ettil		kinase	264488, 264600, 264602, 264764, 264636
417	20153707 (833 834)	Novel Protein sim Chark	137 FOOT OF COLUMN		0000
	(000, 004)	NOVEL FOREIT STATEMENT STA	Contains protein domain (P+00145) - C-5 cytosine-specific DNA methylase		264605
- 1	94125841 (835, 836)			UNCLASSIFIED	264689, 264693
419	95314273 (837, 838)			collagen	264908, 264910, 264764, 264639
- 1	37036349 (839, 840)	Novel Protein sim. GBank gij3261659jembjCAB03751j - (281368) hypothetical protein Rv2419c [Mycobacterium tuberculosis]	Contains protein domain (PF00300) - phosphatase Phosphoglycerate mutase family	phosphatase	264769
421	95292942 (841, 842)	Novel Protein sim. GBank gil 2916942 lemb CAA 17580 - (ALO21999) hypothetical protein Rv0981 [Mycobacterium tuberculosis]	Contains protein domain (PF00072) - phosphatase Response regulator receiver domain	phosphatase	264906, 264600, 264601, 264603, 264604, 264760, 264769
422	79471293 (843, 844)	Novel Protein sim. GBank gi[231752[sp]Q00767[CH61_STRAL - 60 KD CHAPERONIN 1 (PROTEIN CPN60 1) (GROEL PROTEIN 1) (HSP58)	Contains protein domain (PF00118) - eph STRAL - 60 KD CHAPERONIN TCP-1/cpn60 chaperonin family DEL PROTEIN 1) (HSP58)	ерћ	22278996, 264682, 18108376, 18108387
423	79604948 (845, 846)			UNCLASSIFIED	264509

,,,,	70000				
Ş	(041, 040)			struct	265019
425	80431450 (849, 850)	Novel Protein sim. GBank gil 1703701 bbs 178462 -	Contains protein domain (PF00225) - struct	struct	264909, 265007, 55811386, 264768,
		KRP5=kinesin-related protein [rats, testes, Peptide Partial, 167 aa]	Kinesin motor domain		55810764
456	80064522 (851, 852)				264605, 264559
427	80057232 (853, 854)	Novel Protein sim. GBank gi 231829 sp P29929 COBN PSEDE - COBN PROTEIN		UNCLASSIFIED	264603, 264636
428	79487798 (855, 856)			UNCLASSIFIED	264683
429	80091252 (857, 858)	Novel Protein sim. GBank gil81286[pir] S22697 - extensin -			35696423, 35695763, 35695855, 265017,
430	80504102 (850 BED)	Mount Destrict (in agricult)			264564, 264762
3	0000,000,000)	1784395) hynothetical protein Ryness (Mycobacterium		reductase	264508, 264905, 264509, 264908, 264909, 26508, 264600, 264687, 264700, 264000
		[tuberculosis]			zobuus, zo4600, zo4687, zo4769, zo4689, 264636, z64638, 18108385, z64486
431	20624249 (861, 862)				264566
432	16525372 (863, 864)				265020
433	81494303 (865, 866)	Novel Protein sim. GBank gij3123552 emb CAA18609 -		UNCLASSIFIED	264907 264908 264909 264910 264592
		(AL022578) dJ393P12.2 (hypothetical Proline-rich protein			264595, 264758, 264604, 264760, 264762.
		KIAA0269 LIKE) [Homo sapiens]			264763, 264636, 264637, 22279002
434	94326323 (867, 868)	Novel Protein sim. GBank	Contains protein domain (PF00169) -	UNCLASSIFIED	55812038, 56182181, 56181562, 29331828.
		gil2495272lsp Q99626 CDX2_HUMAN - HOMEOBOX	PH domain		35696052, 55810764, 55811576, 65274791,
		PROTEIN CDX-2 (CAUDAL-TYPE HOMEOBOX PROTEIN			35695855, 60432113, 55811150, 264636.
		2) (CDX-3)			264766
435	80502738 (869, 870)	Novel Protein sim. GBank		transport	264595, 264769
		gij114105jspjP08532jARAH_ECOLI - L-ARABINOSE		•	
إ		TRANSPORT SYSTEM PERMEASE PROTEIN ARAH			
8	41085953 (871, 872)			UNCLASSIFIED	265020, 22279002
437	11399291 (873, 874)			UNCLASSIFIED	264593
2	11773835 (875, 876)			UNCLASSIFIED	264686
439	80019495 (877, 878)	Novel Protein sim. GBank gi 3242702 (AC003040) -		Γ	264905 264600 264602 264604
		hypothetical protein [Arabidopsis thaliana]			
4	79841062 (879, 880)		Contains protein domain (PF00004) -	ATPase_associated	Contains protein domain (PF00004) - ATPase_associated 35696052, 264905, 264908, 264909, 265011,
_			ATPases associated with various	1	35696423
		(AAA), Score=268.1, E-value=3.7e-77, N=1 (Caenorhabdilis cellular activities (AAA)	cellular activities (AAA)		
;	200000000000000000000000000000000000000	elegans			
<u> </u>	(20030030) (001, 002)	Novel Protein sim. GBank nij5639946jahjAAD45904 1/AF16132 - /AF161328) histidiaa			264605
		kinase CstS [Corynebacterium diphtheriae]			
442	85281058 (883, 884)	Novel Protein sim. GBank gil1184790 (U46068) - von Ebner		UNCLASSIFIED	29331830, 264909
-		minor salivary gland protein [Mus musculus]			
1	82456427 (885, 886)	Novel Protein sim. GBank gij5689893 emb CAB52056.1 -		UNCLASSIFIED	35696052, 264508, 264906, 264512, 264604,
		(AL109732) putative ATP-binding RNA helicase			264762, 264769, 264689, 264636
444	11305807 /807 0001	Name Destriction of Control of Topological Control			
<u> </u>	(000', 000)	Novel Protein sim. GBank gij1/83249[dbj BAA11726] -		UNCLASSIFIED	264591
		(coooce) inclinations to climate-sociam symport (citrate transporters): hypothetical (Bacillus subtilis)			

445	79552709 (889, 890)			UNCLASSIFIED	264693
446	79810937 (891, 892)	Novel Protein sim. GBank gi 5531272 emb CAB50897.1 - [AJ243800] WSC4 homologue [Kluyveromyces lactis]			264509
447	80438888 (893, 894)	Novel Protein sim. GBank gi[538413 (L36315) - zinc finger protein [Mus musculus]	Contains protein domain (PF00096) - transcriptfactor Zinc finger, C2H2 type	transcriptfactor	264768, 55811576
448	80238110 (895, 896)	Novel Protein sim. GBank gi 1542914 emb CAB02185 - (280108) fmt (Mycobacterium tuberculosis)	Contains protein domain (PF00551) - dehydrogenase Formyl transferase		264508, 264600, 264603, 264605, 264682, 264769, 18108362, 264634, 18108387
449	20460634 (897, 898)	Novel Protein sim. GBank gi[118794 sp P10443 DP3A_ECOL] - DNA POLYMERASE III, ALPHA CHAIN		polymerase	264605, 264559
450	94631210 (899, 900)	Novel Protein sim. GBank gil4589506 dbj BAA76775.1 -	Contains protein domain (PF00481) - phosphatase	phosphatase	65274572, 22278998, 29331824, 29331826, 254006, 254010, 254503
		(Abox 5 146) Niversi profein inomo sapiensi	Profein phosphatase 2C		254905, 254910, 254592, 52545317, 255017, 21906767, 55811957, 56526486, 22279002
451	21433609 (901, 902)			UNCLASSIFIED	264486
452	10267276 (903, 904)				264692
453	52560096 (905, 906)	Novel Protein sim. GBank gi 2650614 (AE001104) - conserved hypothetical protein [Archaeoglobus fulgidus]		UNCLASSIFIED	264907, 264600
454	39523922 (907, 908)	Novel Protein sim. GBank		transferase	264603
		gij2493000 sp Q09450 SCOT_CAEEL - PROBABLE			
		SUCCINYL-COA:3-KETOACID-COENZYME A			
		TRANSFERASE PRECURSOR (3-0X0ACID COA-			
		TRANSFERASE)			
455	13089692 (909, 910)			UNCLASSIFIED	264687
456	79563081 (911, 912)			UNCLASSIFIED	264691
457	79831273 (913, 914)	Novel Protein sim. GBank gij4466699jemb CAB38153.1 - (AL035591) putative integral membrane export protein Streptomyces coelicolor			264905
458	79581227 (915, 916)	Novel Protein sim. GBank gij3411053 (AF034863) -	Contains protein domain (PF00595) - kinase	kinase	55812038, 265010, 265018, 264681
	,	synaptic scaffolding molecule [Rattus norvegicus]	PDZ domain (Also known as DHR or GLGF).		
459	80567359 (917, 918)	Novel Protein sim. GBank	Contains protein domain (PF00130) - kinase	kinase	22278997, 264259, 29331826, 265018,
		gl 4506075 ref NP_002733.1 pPRKC - protein kinase G, mu	Phorbol esters/diacylglycerol binding domain (C1 domain)		264448, 264369, 21906765, 35696423
460	79245890 (919, 920)	Novel Protein sim. GBank		UNCLASSIFIED	264906
_		gi 113158 sp P25516 ACO1_ECOLI - ACONITATE HYDRATASE 1 (CITRATE HYDRO-LYASE 1) (ACONITASE 1)			
461	95287618 (921, 922)	Novel Protein sim. GBank		synthase	264602, 264605, 264768, 264769, 265021,
		gi[1168574 sp P42464 ATPB_CORGL - ATP SYNTHASE BETA CHAIN			33657023, 264559

Г	1,000 000,001000000	1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -	Castains protein domain (DE00304) II (NC) ASSIEIED		264007
701	(***) 6000000	WOVER TUGE!! SIII. COOKING TO THE CONTROL OF THE CO			
463	79796417 (925, 926)	Novel Protein sim. GBank gil854065[emb CAA58337] - (X83413) U88 [Human herpesvirus 6]		UNCLASSIFIED	264905, 264906, 264908, 264909, 264910, 264591, 264591, 264595, 265011, 264632, 264635,
;	00040464 (007 000)			INC. ASSIEIED	264636, 264637, 264638, 264639 264634
į į	026340131 (927, 920)	1 CONTRACTOR CONTRACTO	October States demois (DE00649)	discoult in the second	25E017 21006784 265020
465	83005730 (929, 930)	Novet Protein sim. GBank gijb689776 emplCAB52137.1 - (AJ242832) calpain [Homo sapiens]	Contains protein domain (PF00046) - Camepsin Calpain family cysteine protease	camepsin	203011, 21306704, 203020
466	20460645 (931, 932)	Novel Protein sim. GBank gil 1806 175 emb[CAB06470] - (Z84395) rpsC [Mycobacterium tuberculosis]	Contains protein domain (PF00417) - ribosomalprot Ribosomal protein S3, N-terminal domain.	ribosomalprot	264605, 264559
467	80409035 (933, 934)	Novel Protein sim. GBank		UNCLASSIFIED	264764
		gijo46703 spip-36949 r656_BACSU - U-RIBUSE-BINDING PROTEIN PRECURSOR			
468	52562208 (935, 936)			UNCLASSIFIED	264692
469	19520527 (937, 938)	Novel Protein sim. GBank gi[2114024 emb CAB08957 - (295558) grcC1 [Mycobacterium tuberculosis]		UNCLASSIFIED	264488
470	80502756 (939, 940)	Novel Protein sim. GBank gil2909459 emb CAA17347 - (AL021929) cobQ [Mycobacterium tuberculosis]		synthase	264602, 264769
471	17937351 (941, 942)	Novel Protein sim. GBank gij114921 sp P17447 BETT_ECOLJ - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN		transport	265019
472	80047458 (943, 944)				264596, 264685, 264557
473	20558793 (945, 946)	Novel Protein sim. GBank gil862343 (L10908) - Gcap1 gene product [Mus musculus]		UNCLASSIFIED	264369
474	80593365 (947, 948)			UNCLASSIFIED	22278997, 264692, 264288
475	82454665 (949, 950)			UNCLASSIFIED	264907, 264908, 264511, 265009, 264762, 264448, 264636, 264638
476	94143857 (951, 952)	Novel Protein sim. GBank gi[5453656 ref NP_006329.1 pGAC1 - glioma amplified on chromosome 1 protein (teucine-rich)	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat	glycoprotein	65274572, 60432049, 264259, 264508, 52644045, 55812038, 264758, 265011, 264288, 264686, 52644229, 65274791, 264638, 264586
477	79175833 (953, 954)			UNCLASSIFIED	264636
478	79633483 (955, 856)			UNCLASSIFIED	264690, 264693
479	80189746 (957, 958)			collagen	264686, 35695855, 265008, 264631, 264910, 264632, 264638, 265018, 264369, 264909
98 80	79390729 (959, 960)	Novel Protein sim. GBank gi[1127551 (U18939) - orf2 [Battrachocottus baikalensis]		rnapolymerase	264369
481	79624578 (961, 962)			UNCLASSIFIED	264693
482	83050611 (963, 964)	Novel Protein sim. GBank gil4063042 (AF088065) - GP900; mucin-like glycoprotein [Cryptosporidium parvum]		UNCLASSIFIED	264909, 264686, 264768, 264693, 55811576, 56182323, 18108385

483	20293306 (965, 966)	Novel Protein sim. GBank gil2104303 emb CAB08632 - (295387) hypothetical protein Rv2610c [Mycobacterium	Contains protein domain (PF00534) - Glycosyl transferases group 1		264600
		tuberculosis]			
48 4	11618046 (967, 968)	Novel Protein sim. GBank gij3450883 (AF083334) - fibroin [Antheraea pemyi]		UNCLASSIFIED	264594
485	80191234 (969, 970)			UNCLASSIFIED	264369, 21906765, 22279000, 22279002
	80059042 (971, 972)	Novel Protein sim. GBank gij5042272 emb CAB44526.1 - (AL078618) nuoF, NADH dehydrogenase subunit [Streptomyces coelicolor]		dehydrogenase	264604
487	11813339 (973, 974)				264638
488	91222383 (975, 976)	Novel Protein sim. GBank gi[5724778 gb AAC53522.2 -	Contains protein domain (PF00620) -		264686, 66714117, 264768, 18108385,
		(AF012273) mo-type GTPase-activating protein rhoGAPX-1 Mus musculus]	RhoGAP domain		55811576, 265006, 265008, 265009, 265019, 22279002, 264259, 18108370, 264907,
-					264764, 56182323, 264288, 264693
489	10867710 (977, 978)	Novel Protein sim. GBank gij3882223]dbj BAA34471.1 - (AB018294) KIAA0751 protein [Homo sapiens]		kinase	264639
490	95361124 (979, 980)	Novel Protein sim. GBank gij82091 piri A25494 -		collagen	22278996, 29331822, 29331828, 264107,
		hydroxyproline-rich glycoprotein - tomato (fragment)			264909, 264110, 265009, 264592, 264593,
					60433356, 264288, 264693, 263974, 263976, 20281071, 60432113
491	80496412 (981, 982)	Novel Protein sim. GBank gij2894206 embjCAA17072 - (AL021840) hypothetical protein Rv3258c (Mycobacterium Iuberculosis)		UNCLASSIFIED	264769
492	87421264 (983, 984)				264600
493	(11692942 (985, 986)			UNCLASSIFIED	264638
484	87726604 (987, 988)	Novel Protein sim. GBank gil5262605jembjCAB45743.11 -		UNCLASSIFIED	264489, 35696286, 60432289, 29331828,
	•	(AL080150) hypothetical protein [Homo sapiens]			35696052, 264509, 264905, 264906, 264907,
					264908, 264909, 264510, 264511, 265009,
					264910, 33657402, 264762, 264764, 264768, 3643769, 364368, 364468, 364686, 364686, 364686, 364686, 364686, 3646686, 364668, 364668, 364668, 3646680, 3646680, 3646680, 3646680, 364680, 364680, 364680, 364680, 364680, 364680, 364680, 364680
					204103, 204000, 21900103, 21900103, 36606017 366030 366603 33657100
					33033317, 203020, 204033, 33037103, 2564620 25606422 25605855 264634
					264638
495	80028599 (989, 990)	Novel Protein sim. GBank gi[2791517]emb CAA16054] -	Contains protein domain (PF00005) - transport	transport	264602, 264682, 264638
		[uberculosis]			
496	78985624 (991, 992)	Novel Protein sim. GBank gi[230281]pdb[1R69] - 434	Contains protein domain (PF01381) -		264601, 265021
		Repressor (Amino-Terminal Domain) (R1-69)	Helix-tum-helix		
497	78949661 (993, 994)	Novel Protein sim. GBank		oxidase	265006
		gij129736 spip28225 PDXH_ECOLI - PYRIDOXAMINE 5- PHOSPHATE OXIDASE (PNP/PMP OXIDASE)			

498	88095488 (995 996)	Novel Protein sim GBank pil1145789 (1141662) - neuroligio	Contains protein domain (PE00135) -	peterses	26425G 20331826 35606052 264508
		2 [Rattus norvegicus]	Carboxylesterases		264509, 264905, 264906, 264907, 264908,
					264909, 264510, 264511, 265009, 264910,
					264591, 33637402, 264756, 265010, 265011, beareon peacon poacon
					20400U, 2040UI, 2040US, 20400S, 204704. 1264766 264767 264768 264687 264760
					204/00, 204/01, 204/00, 204/01, 204/03, 21000267, 204/03,
					21900101, 33031023, 204033, 204020, 364630, 36606423, 264630, 364632, 364634
					204023, 30030423, 204030, 204032, 204034, 264635, 264637, 264638, 264558, 264639
					18108385 264563 264564 264565 264565
					18106363, £04363, £04364, £04363, £04369,
499	20438222 (997, 998)	Novel Protein sim. GBank gil97480 pir S19739 - integral		UNCLASSIFIED	264605
		membrane protein - Rhodobacter capsulatus			
200	11076810 (999, 1000)				264605
501	13418034 (1001, 1002) Novel Protein sim. G	Novel Protein sim. GBank gil5708250 emb CAB52363.1 -		UNCLASSIFIED	264688
		(AL 109747) putative integral membrane protein			
		(Streptornyces coelicolor A3(2))			
205	80021176 (1003, 1004)	80021176 (1003, 1004) Novel Protein sim. CBank gil4466578 emb CAB38132.1 - (AB38132.1) emplores (Streptomyces CAB3813) glucose-6-phosphate isomerase (Streptomyces CAB381) glucose-6-phosphate isomerase (Streptomyces CAB38181) glucose-6-phosphate isomerase (Streptomyces CAB3818181) glucose-6-phosphate isomerase (Streptomyces CAB381818181818181818181818181818181818181	Contains protein domain (PF00342) - isomerase Phosphoglucose isomerase	isomerase	22278996, 265011, 264602, 264605, 264635
503	20264483 (1005 100B)			JINO ACCIETED	20,456.4
3 2	10887324 (1007 1008)			ONCLASSIFIED	100,000
5	10001321 (1007, 1000)			UNCLASSIFIED	784687
ရှိ	95003068 (1009, 1010)			UNCLASSIFIED	264369
206	16454292 (1011, 1012) Novel Protein sim. G	Novel Profein sim. GBank	Contains protein domain (PF00036) - struct	struct	265010
1		gilanggonalsplanggalcarm IETFT - CALMODOLIN	Er nand		
20,	20451598 (1013, 1014)	20451598 (1013, 1014) Novel Protein sim. GBank		UNCLASSIFIED	264604
		[gi[2501069]sp[Q46127]SYW_CLOLO - TRYPTOPHANYL- Tena SYNTHETASE (TEXPEDIAN) TENA 10005			
		(TRPRS)			
208	79841424 (1015, 1016)	79841424 (1015, 1016) Novel Protein sim. GBank		UNCLASSIFIED	264908
		gi 466068 sp P34618 YO82			
909	11776386 (1017, 1018)				264638
510	83373465 (1019, 1020)			UNCLASSIFIED	264687, 264639
511	16525578 (1021, 1022)				265007
512	20399484 (1023, 1024) Novel Protein sim. G	Novel Protein sim. GBank		UNCLASSIFIED	264565
		gi 2497419 sp P55635 Y4RB_RHISN - PUTATIVE INTEGRASE/RECOMBINASE Y4RB			
513	79457404 (1025, 1026)	79457404 (1025, 1026) Novel Protein sim. GBank gij1276897 (U41809) - cyclin J	Contains protein domain (PF00134) - cyclin	cyclin	264683, 264689, 35696423, 264639
		[Orosophita melanogaster]	Cyclin		
514	(79813805 (1027, 1028)	79813805 (1027, 1028) Novel Protein sim. GBank gil1184790 (U46068) - von Ebner		UNCLASSIFIED	29331830, 264909
		minor salivary gland protein [Mus musculus]			
2	79462591 (1029, 1030)				22278999, 264690
916	9862020 (1031, 1032)	Novel Protein sim. GBank gi 2127400 pir S65770		amylase	264910
		mattooligosyttrehatose trehalohydrolase - Arthrobacter sp. (strain O36)			

		T		_	,	,		_	_			_	-	,	_	
265018, 264605, 264764, 264766, 264687	264691, 264565	264487	52644507, 22278997, 22278998, 60432049, 264259, 52645080, 29331824, 66714117, 60424269, 29331826, 35696052, 264905, 29331830, 66712502, 264511, 265007, 264591, 60432229, 33657402, 60433438, 21906754, 33109954, 52644296, 8716859, 265017, 265018, 264604, 265019, 264681, 264681, 26448, 264589, 264685, 21906756, 21906756, 21906756, 21906766, 2960767, 21906759, 265021, 60170915, 33657023, 246624, 3569763, 18108370, 264629, 52644332, 56182323, 264639, 83373044, 18108385, 56526486, 60432113	264769	264605	264905, 264768	264629	264092, 264596, 265011	264907	264758	264769	56182575, 265017, 265018	265019	264687	264555, 264556, 264557, 264558, 18108385	264488, 264490, 264259, 264592, 264760,
isomerase		UNCLASSIFIED	dna_rna_bind		transport	kinase	UNCLASSIFIED	ATPase_associated		UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	hydrolase	UNCLASSIFIED	UNCLASSIFIED	
			Contains protein domain (PF01388) - dna_rna_bind ARID DNA binding domain			Contains protein domain (PF00294) - kinase pfkB family carbohydrate kinase										Contains protein domain (PF00622) - UNCLASSIFIED SPRY domain
95292994 (1033, 1034) Novel Protein sim. GBank gil2983605 (AE000725) - ribose 5	phosphate isomerase B [Aquifex aeolicus]	Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 (Human herpesvirus 6]	91677886 (1037, 1038) Novel Protein sim. GBank gijs689365jdbjjBAA83073.11 - (AB024075) B120 [Homo sapiens] 78869188 (1030, 1040).		110/6821 (1041, 1042) Novel Protein sim. GBank gij1169126 sp P46839 CTPA_MYCLE - CATION- TRANSPORTING P-TYPE ATPASE A	80435060 (1043, 1044) Novel Protein sim. GBank gi 1172869 sp P44331 RBSK_HAEIN - RIBOKINASE	18356013 (1045, 1046) Novel Protein sim. GBank gi 2132243 pir S61028 - hypothetical protein YPL236c - yeast (Saccharomyces cerevisiae)	80261805 (1047, 1048) Novel Protein sim. GBank gil4033608 dbj BAA35136 - (AB012308) B2HC [Anthocidaris crassispina]		Novel Protein sim. GBank gil4106610 emb CAA21365 - (AL031866) ORF42. len=386 aa , similarity to an aminotransferase, in P95957 Sulfolobus solfataricus, (401 aa), 31.% identity in 393 aa overlap, Fasta scores: opt:468, E(): 8.56-24, in Q64602 R. norvegicus, (425 aa), 28.6% identi.	80504729 (1053, 1054)		17936810 (1057, 1058) Novel Protein sim, GBank gij731088 sp P24215 UXUA_ECOLI - MANNONATE DEHYDRATASE (D-MANNONATE HYDROLASE)	Novel Protein sim. GBank gil42144 jemb CAA25200] - (X00513) NusA protein (nusA) [Escherichia coli]		Bank gij5262640jembjCAB45758.1j -
95292994 (1033, 1034)	10000	6491831 (1035, 1036)	91677886 (1037, 1038) 79869488 (1030, 1040)	1909100 (1038, 1040)	110/6821 (1041, 1042)	80435060 (1043, 1044)	18356013 (1045, 1046)	80261805 (1047, 1048)	79610046 (1049, 1050)	36827630 (1051, 1052) Novel Protein sim. Gl (AL031866) ORF42, aminotransferase, in aa), 33.1% identity in opt:468, E(): 8.56-24, 28.6% identi.	80504729 (1053, 1054)	65484134 (1055, 1056)	17936810 (1057, 1058)	10887336 (1059, 1060)	80226576 (1061, 1062)	90933444 (1063, 1064) Novel Protein sim. GI (AL080170) hypothet
217	- 1	ρ Ω	918		176	225	223	524	222	226	527	528	529	230	231	232

533	87761531 (1065 1066)	87761531 (1065 1068) Novel Protein sim GBank			PEAGNT PEAGNG PEATER REGEGIT PEAER
<u> </u>	(2001)	gil4883636[gb]AAD31593.1[AF11229 - (AF112299) integral			264555
		inner nuclear membrane protein MAN1 (Homo sapiens)			
534	82368264 (1067, 1068)	82368264 (1067, 1068) Novel Protein sim. GBank gi 2995352 emb CAA04606.1 -		UNCLASSIFIED	264905, 265011, 264601, 264602, 264605,
		(AJ001206) pep1 [Streptomyces coelicolor]			264762, 264768, 265020, 264693, 264636
235	79641850 (1069, 1070)	79641850 (1069, 1070) Novel Protein sim. GBank gij3878636jemb CAA88953j - (Z49128) similar to cAMP-dependant protein kinase; cDNA EST EMBL:T00719 comes from this gene; cDNA EST yk46548.3 comes from this gene; cDNA EST	Contains protein domain (PF00069) - ATPase_associated 264906 Eukaryotic protein kinase domain	ATPase_associated	264906
		comes from this gene; cDNA EST yk492f4.3 comes from this gene; cDNA EST $\gamma_{\rm eff}$			
536	79907207 (1071, 1072)	79907207 (1071, 1072) Novel Protein sim. GBank gi[2495628 sp P55757 YOHL_SERMA - HYPOTHETICAL		reductase	18108376, 264905, 264906, 264907, 264909
		10.1 KD PKO LEIN IN BIOA S'REGION			
237	94147448 (1073, 1074)				265008, 264605, 65274791
238	87821963 (1075, 1076)	87821963 (1075, 1076) Novel Protein sim. GBank	Contains protein domain (PF00595) - collagen	collagen	29331822, 29331824, 29331825, 29331826,
		gif134920jspjP21997 SSGP_VOLCA - SULFATED	PDZ domain (Also known as DHR or	_	29331827, 264908, 52644045, 33657402,
		SURFACE GLYCOPROTEIN 185 (SSG 185)	GLGF).		265017, 264762, 264683, 264288, 264685, 21906765, 35695763, 264558, 60170394, 264559, 22279002
539	28396269 (1077, 1078)	28396269 (1077, 1078) Novel Protein sim. GBank		histone	264602, 265019
		gji2498433jspjQ12341JHAT1_YEAST - HISTONE ACETY TRANSEERASE			
540	79637077 (1079, 1080)				264693
17	8776226R (1081 1082)	87762268 (1081 1082) Novel Protein sim GBant Gil 3882241 dhill 4034480 11	Contains protein domain (PE00096)	transcriptfactor	18108394 22278997 22278998 264259
<u> </u>		(AB018303) KIAA0760 protein [Homo sapiens]	Zinc finger, C2H2 type	מפוסקייספיסן	264112, 265009, 33657402, 55512038, 52646317, 265017, 21906765, 264693, 55811576, 264655, 56526486, 264566
542	95295836 (1083, 1084)	95295836 (1083, 1084) Novel Protein sim. GBank gij5042272 embJCAB44526.1 - (AL078618) nuoF, NADH dehydrogenase subunit Streptomyces coelicolor		dehydrogenase	264910, 265018, 264689, 264638, 264486
543	79796290 (1085, 1086)	-		UNCLASSIFIED	264602, 264908
4	20437191 (1087, 1088)	20437191 (1087, 1088) Novel Protein sim. GBank gil2791398 emb CA415994 - (AL021184) hypothetical protein Rv1464 [Mycobacterium tuberculosis]		UNCLASSIFIED	264605
545	80434504 (1089, 1090)				264768, 264634, 264907, 264592, 264909
546	80249016 (1091, 1092)	80249016 (1091, 1092) Novel Protein sim. GBank gil4887211[gb AAD32237.1 AF14744 - (AF147449) penicillin			264600, 264602, 21906765
		binding protein 1B [Pseudomonas aeruginosa]			
2	11077563 (1093, 1094)	11077563 (1093, 1094) Novel Protein sim. GBank		rnapolymerase	264604
		gij 350855 Sp +19176 R+OC FSEFU - UNA-UIKECI EU RNA FOLYMERASE BETA: CHAIN (TRANSCRIPTASE RFTA: CHAIN) (PNA POL YMEDASE BETA: SLIBI INITY			
548	82114936 (1095, 1096)	82114936 (1095, 1096) Novel Protein sim. GBank gij2330021 (AF019250) - kinesin-		UNCLASSIFIED	264488, 264905, 264910, 264760, 264693,
		related protein; KRP; Costal2 [Drosophita melanogaster]			264639, 264563, 264564

27 6	95421904 (1097, 1098)	95421904 (1097, 1098) Novel Protein sim. GBank gild337460 gblAAD18133] - (AF056195) neuroblastoma-amplified protein [Homo sapiens]		UNCLASSIFIED	264488, 65274572, 18108398, 22278995, 22278996, 22278997, 22278998, 22278999, 226429, 22278998, 22278999, 264259, 2331824, 66714117, 29331825, 2643331826, 35696052, 265007, 265008, 264910, 264992, 33657402, 33109954, 265017, 265018, 26409, 18108351, 26448, 264688, 21906765, 21906766, 21906767, 21906768, 21906765, 21906766, 21906767, 23906768, 21906769, 264691, 2357023, 264692, 264693, 65274620, 26245129, 33657109, 27486261, 27486262, 264634, 33657349, 55811576, 18108387, 61432113, 222740102
220	10886616 (1099, 1100)				264688
551	80439990 (1101, 1102)	80439990 (1101, 1102) Novel Protein sim. GBank gi 3122893 sp P94985 SYFB_MYCTU - PHENYLALANYL- TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE TRNA LIGASE BETA CHAIN) (PHERS)		UNCLASSIFIED	264908, 264909, 264768
222	94672870 (1103, 1104)	94672870 (1103, 1104)		UNCLASSIFIED	264689, 264639, 264563
223	80106002 (1105, 1106)	Novel Protein sim. GBank gil552087 (M33753) - crumbs protein [Drosophila melanogaster]	Contains protein domain (PF00008) - glycoprotein EGF-like domain	glycoprotein	55811957, 264628
400	79618379 (1107, 1108)	786 18379 (1107, 1108) Novel Protein sim. GBank gi 5019771 gb AAD37857.11AF13326 - (AF133263) histidine protein kinase-response regulator hybrid protein CvgSY [Pseudomonas syringae pv. syringae]		kinase	264906
555	78996347 (1109, 1110)	A_SALTY - PTS SYSTEM. COMPONENT (EIIA-GLC) IA COMPONENT) SE ENZYME II, A COMPONENT)	Contains protein domain (PF00358) - transport phosphoenolpyruvate-dependent sugar phosphotransferase system, EIIA 1	transport	264762
226	20457127 (1111, 1112)	20457127 (†111, 1112) Novel Protein sim. GBank gij3914014 sp P96380 MFD_MYCTU - TRANSCRIPTION- REPAIR COUPLING FACTOR (TRCF)		transcriptfactor	264508, 264605, 264559
/66	19523405 (1113, 1114)	19523405 (1113, 1114) Novel Protein sim. GBank gij5042273jembjCAB44527.1j - (AL078618) nuoE, NADH dehydrogenase subunit [Streptomyces coelicolor]		dehydrogenase	264488
8888	20724429 (1115, 1116)	20724429 (1115, 1116) Novel Protein sim. GBank gil170933lsplP45331METE. HAEIN - 5- METHYTETRAHYDROPTEROYLTRGLUTAMATE HOMONY TRAINSTERASE (METHIONINE SYNTHASE, VITAMIN-B12 INDEPENDENT ISOZYME) (COBALAMIN-INDEPENDENT METHIONINE SYNTHASE)		UNCLASSIFIED	264602
9229	80084353 (1117, 1118)	80084353 (1117, 1118) Novel Protein sim. GBank gil4980567[gb AAD35173.1 AE00169 - (AE001694) iron(III) ABC transporter, permease protein [Thermotoga maritima]		UNCLASSIFIED	264634

١					
280	80059417 (1159, 1160) 	(09)			22278999, 35696052, 264555, 264556, 264558
581	79230833 (1161, 1162)	(25)		UNCLASSIFIED	265008, 264564
282	80049617 (1163, 11	gij3243131 (AF045777) - titin	Contains protein domain (PF00047) - struct	struct	265021, 264555, 264557
6	20, 10, 10, 10, 10, 10, 10, 10, 10, 10, 1		Immunoglobulin domain		
28	79321392 (1165, 11	/8321392 (1165, 1166) Novel Protein sim. GBank gil2601162 splf7726 f747ECOLI - HYPOTHETICAL 400 KD BDOTEN IN ABDA CYCE INTERPEDITOR		transport	264594
584	79845024 (1167, 1168)	13:0 15:01		INC ASSISTED	26449B 264006 264766 264687 36606422
585	79581454 (1169, 11	79581454 (1169, 1170) Novel Protein sim. GBank gil38822211dhilBAA34470 11.		INCLASSIFIED	265018 254684 24006750
		(AB018293) KIAA0750 protein (Homo sapiens)		ONCEASSIFIED	Z02016, Z04004, Z1800109
286	38277486 (1171, 1172)	72)		UNCLASSIFIED	264908, 265007
287	80497359 (1173, 11	80497359 (1173, 1174) Novel Protein sim. GBank gil4467250 emb CAB37575 -		hydrolase	264600, 264602, 264605, 264769, 264690.
		(AL035569) probable Glu-tRNA Gln amidotransferase subunit [Streptomyces coelicolor]		•	264557
588	79557239 (1175, 11	_		UNCLASSIFIED	265020, 264692
289	79805828 (1177, 1178)			UNCLASSIFIED	22278996, 264907, 264909, 264510, 265009, 265010, 264687, 264769, 35695917,
					18108376, 264634, 264636, 264638
8	79815629 (1179, 1180)	(90)		UNCLASSIFIED	264906, 264909
591	10313540 (1181, 11	10313540 (1181, 1182) Novel Protein sim. GBank gi 2143293 emb CAB09390 - (Z95972) rpoB [Mycobacterium tuberculosis]		mapolymerase	264691
295	13889767 (1183, 1184)	84)		MHC	263972
593	82348699 (1185, 11	82348699 (1185, 1186) Novel Protein sim. GBank gil4511983 gbJAAD21543.1 -		dehydrogenase	264511, 264762, 264769, 264486
	,	(AF088896) electrotransfer ubiquinone oxidoreductase [Zymomonas mobilis]			
594	20212392 (1187, 11	20212392 (1187, 1188) Novel Protein sim. GBank gi[1272368 (U51896) - LígE [Vibrio parahaemolyticus]		UNCLASSIFIED	264605
595	10064064 (1189, 11	10064064 (1189, 1190) Novel Protein sim. GBank			264769
		gil131490lspIP20966IPTFB_ECOLI - PTS SYSTEM. FRUCTOSE-SPECIFIC IIBC COMPONENT (EIBC-FRU)			
		(FRUCTOSE-PERMEASE IIBC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, BC			
296	13085170 (1191, 1192)	92)		UNCLASSIFIED	264636
265	80259003 (1193, 1194)	94)		UNCLASSIFIED	264592
298	94140216 (1195, 1196)	(96)		UNCLASSIFIED	264758, 55810764, 264555, 264556, 264637, 83373044
669	20385137 (1197, 11	20385137 (1197, 1198) Novel Protein sim. GBank gil125329lspP04951 KDSB_ECOLI - 3.DEOXY-MANNO- OCTULOSONATE CYTIDYLYLTRANSFERASE (CMP-KDO SYNTHETASE] (CMP-2.KETO-3.DEOXYOCTULOSONIC ACID SYNTHETASE, (CKR)		UNCLASSIFIED	264603
8	10357663 (1199, 1200)				264906
9	79610404 (1201 12	79610404 (1201, 1202) Novel Protein sim, GRank oil 2127414 Inivitie 600 Ra		I INC. ACCIETED	20000
		hypothetical protein 2 - Corynebacterium glutamicum		UNCLASSIFIED	264510

265007	264595	264758	264605	264764	264508, 264906, 85658542, 264687, 264687, 264689, 264534, 18108376, 35696423, 264536, 264555, 264638	264682	264605	264692	264508, 264905, 264907, 264908, 264909, 264511, 264910, 264758, 264604, 264684,	264766, 284689, 264692, 264628, 264635, 264636, 264637, 264558	264600, 264601, 264604, 264769, 264558,	264565	264369	18108372, 264563	264600, 264602, 264629	264605		52645156, 21906765, 35696423, 21906768,	21906769, 22278994, 35696286, 22278996, 265020, 265021, 265007, 265008, 264636.	52644150, 33657023, 264692, 264693,	29331822, 29331824, 55812038, 83373044, 56182181, 60424269, 66714117, 29331825.	33657109, 29331826, 33657182, 29331827,	35696052, 29331828, 27486262, 33657349,	56526486, 265018, 265019, 22279002, 264482, 264448, 29331830, 66712502,	264909
kinase	UNCLASSIFIED		UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED				nuclease		UNCLASSIFIED	kinase	synthase	isomerase		tnf					· 		
Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain							Contains protein domain (PF00986) - UNCLASSIFIED PAS domain				Contains protein domain (PF01443) - nuclease	Viral (Superfamily 1) RNA helicase						Contains protein domain (PF00641) -	Zn-finger in Ran binding protein and others						
79250602 (1203, 1204) Novel Protein sim. GBank gi[3522961[gb]AAC34243.1] - (AC004411) putative pto kinase [Arabidopsis thaliana]		(8)	20436657 (1209, 1210) Novei Protein sim. GBank gij1175322 sp P44917 Y883_HAEIN - HYPOTHETICAL PROTEIN H10883	80334582 (1211, 1212) Novel Protein sim. GBank gi 5020264 gb AAD38043.1 AF15136 - (AF151363) Cdc42 GTPase-activating protein [Mus musculus]	95361506 (1213, 1214) Novel Protein sim. GBank gi 188864 (M74027) - mucin [Homo sapiens]	(9)	80064775 (1217, 1218) Novel Protein sim. GBank gi[2496701 sp P55552 Y4LL_RHISN - HYPOTHETICAL 91.8 KD PROTEIN Y4LL		(22)		95287851 (1223, 1224) Novel Protein sim. GBank gi 1877366 emb CAB07118 -	(292772) recD [Mycobacterium tuberculosis]	H	79969348 (1227, 1228) Novel Protein sim. GBank gij5114231jgb AAD40238.1 AF13670 - (AF136709) histidine kinase YvcG [Staphylococcus aureus]	39586996 (1229, 1230) Novel Protein sim. GBank gil 1339950 dbj BAA12741 - (D85230) large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	20465331 (1231, 1232) Novel Protein sim. GBank gij544367lsplP35673lGALE_ERWAM - UDP-GLUCOSE 4-	EPIMERASE (GALACTOWALDENASE) (UDP- GALACTOSE 4-EPIMERASE)	91227222 (1233, 1234) Novel Protein sim. GBank	gi[2498097]sp[QE0769]TNP3_MOUSE - TUMOR	(PUTATIVE DNA BINDING PROTEIN A20) (ZINC FINGER	PROTEIN A20)				
79250602 (1203, 1204	11466067 (1205, 1206)	81675420 (1207, 1208)	20436657 (1209, 1210	80334582 (1211, 1212	95361506 (1213, 1214	11810888 (1215, 1216)	80064775 (1217, 1218	79629413 (1219, 1220	87586205 (1221, 1222)		95287851 (1223, 1224		7523475 (1225, 1226)	79969348 (1227, 122)	39586996 (1229, 123	20465331 (1231, 123		91227222 (1233, 123							
905	Т	Т	909	909	209	809	609	Г	611		612		613	614	515	616		617		_					

1235, 1236	20632843 (1235, 1236) Novel Protein sim. GBank gi[5459388]emb CAB50746.1 - (AL096839) putative aminotransferase [Streptomyces coelicolor]		isomerase	264603
91227224 (1237, 1238)				56994075, 29331826, 33656970, 265008, 33657402, 33109954, 87166559, 264448, 18108374, 83373044
~	81183143 (1239, 1240) Novel Protein sim. GBank gil464335[sp Q05922[DUS2_MOUSE - DUAL SPECIFICITY PROTEIN PHOSPHATASE 2 (DUAL SPECIFICITY PROTEIN PHOSPHATASE PAC-1)		phosphatase	29146498, 264758, 264369, 29148627
80239251 (1241, 1242)			UNCLASSIFIED	264556 264558 264639
_	20456427 (1243, 1244) Novel Protein sim. GBank gi[2633557 emb CAB13060 - (299110) yidF [Bacillus subtilis]		UNCLASSIFIED	264605
	57710 gb AAB48482 - protein [Rattus norvegicus]	Contains protein domain (PF00054) - Iaminin Laminin G domain	laminin	264906
			cytochrome	264596
	13084619 (1249, 1250) Novel Protein sim. GBank gi[2894252 emb CAA77114.1] - (AL021841) hypothetical protein Rv3342 [Mycobacterium tuberculosis]		UNCLASSIFIED	264668
			UNCLASSIFIED	29331822, 264905, 264908, 33657023, 33657109, 264558
			UNCLASSIFIED	18108394, 264112, 264593, 265022, 264635
:		Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	kinase	264600
I			somerase	264906, 264907
~ 1	39565156 (1259, 1260) Novel Protein sim. GBank gij3236368 (AF064748) - S3-12 [Mus musculus]		UNCLASSIFIED	264490
2.5	20598718 (1261, 1262) Novel Protein sim. GBank gil140687lspIP11666 YGGB_ECOLI - HYPOTHETICAL 30.9 KD PROTEIN IN SBM-FBA INTERGENIC REGION (ORF 4) (F286)			263978
27843890 (1263, 1264)			UNCLASSIFIED	264906, 264600, 264605, 264769, 264689, 264486
80477772 (1265, 1266)			UNCLASSIFIED	264769
17938806 (1267, 1268)				265019
79574506 (1269, 1270)			UNCLASSIFIED	264689
			UNCLASSIFIED	264596, 264762, 264693

637					The second secon	
	82455/96 (1273, 121	82455796 (1273, 1274) Novel Protein sim. GBank gi[2326739 emb CAB10953] - 1798258) rech IMvrohertering in trhetrulogis		nuclease	264906, 264907, 264510, 264511, 264601, 264602, 264603, 264604, 264606, 18108364	
					264762, 264766, 264687, 264769, 264689	
					35695917, 264693, 264634, 264638, 264639	
	:				264559, 18108385	
638	14997457 (1275, 127	14997457 (1275, 1276) Novel Protein sim. GBank gil4678662 emb CAB41074.1 -			264636	_
		(Streptomyces coelicolor)				
629	80204210 (1277, 127	80204210 (1277, 1278) Novel Protein sim. GBank gil4589628 dbjjBAA76836.1 - (AB023209) KIAA0992 protein [Homo caniens]		struct	264112, 263974	_
640	17929579 (1279, 128	17929579 (1279, 1280) Novel Protein sim, GBank qi11432083 (U60981) - homolog Cor	Contains protein domain (PF01466) - Irnapolymerase		265009, 265010	$\overline{}$
	•		cp1 family			
		budding yeast [Arabidopsis thaliana]				
641	79636398 (1281, 1282)				264693	_
642	19898737 (1283, 1284)	34)		UNCLASSIFIED	264565	_
643	81516220 (1285, 1286)	(9)		UNCLASSIFIED	264906, 264908, 264758, 264288, 264632, 264635, 264639, 264564	
644	11751367 (1287, 1288)	(88)		UNCLASSIFIED	264684	
645	95010907 (1289, 1290)	(0)		UNCLASSIFIED	264906, 264762, 264693, 264639, 264559	
646	80069083 (1291, 1292)	32)			264595, 264566	Τ-
647	80257085 (1293, 125	80257085 (1293-1294) Novel Protein sim GBank	Contains profein domain (PE00023) - Itranscriptfactor		264909 264591	1
		111	Ank repeat		201001 201001	_
648	80077428 (1295, 129	80077428 (1295, 1296) Novet Protein sim. GBank gi 1044963 bbs 169646 - protamine [Monodonta turbinata, gonads. Peptide, 106 aa]		UNCLASSIFIED	264600	_
649	80247447 (1297, 1298)	18)		UNCLASSIFIED	263978	_
650	11798316 (1299, 1300)	(00		UNCLASSIFIED	264686	_
651	11776932 (1301, 130	11776932 (1301, 1302) Novel Protein sim. GBank gjl1346916[sp[P12283]PURA_ECOL! -			264602, 264638	
		ADENYLOSUCCINATE SYNTHETASE (IMPASPARTATE LIGASE)				
652	85516704 (1303, 1304)	()4)		UNCLASSIFIED	264905, 264907, 264909, 263978, 264637	_
653	82124947 (1305, 130	82124947 (1305, 1306) Novel Protein sim. GBank		UNCLASSIFIED	22278996, 264510, 264511, 264512, 264593,	_
		gi 1722977 sp Q10638 Y03C_MYCTU - HYPOTHETICAL 82.8 KD PROTEIN CY130.12C			21906754, 264603, 264760, 18108376, 264556	
654	95010589 (1307, 1308)	(80		UNCLASSIFIED	264906, 264595, 264632	$\overline{}$
655	79320692 (1309, 131	79320692 (1309, 1310) Novel Protein sim. GBank 79320692 (1309, 1310) 9i 130327 sp P26647 PLSC_ECOLI - 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (1-AGPACYLTRANSFERASE) (1-AGPAT) (LYSOPHOSPHATIDICACYLTRANSFERASE) (1-AGPAT) ACYLTRANSFERASE) (LPAAT)	Contains protein domain (PF01553) - Itransferase Acyltransferase	transferase	264592	
986	90446730 (1914 191				CONTRACT SOLVED SOLVED	
000	206410/39 (1311, 1312)	(7)			2646UZ, 2645UZ, 264700, 264091	
/20	2061 1010 (1313, 13	[4]		UNCLASSIFIED	264557, 264558	\neg

858	87761915 (1315, 1316) Novel Protein sim. G (AB029001) KIAA107 AB029001 KIAA107	Novel Protein sim. GBank gi 5689493 dbj BAA83030.1 - (AB029001) KIAA1078 protein [Homo sapiens]		UNCLASSIFIED	22278996, 60432049, 29331822, 29331824, 29331828, 265007, 265009, 33657402, 33657084, 265017, 264448, 21906765, 21906766, 263967, 20281149, 18108370,
					18108374, 264482
629	87718663 (1317, 1318)	87718663 (1317, 1318) Novel Protein sim. GBank gi[2137872]pir !48724 - zinc	Contains protein domain (PF00096) - transcriptfactor	transcriptfactor	22278999, 60432049, 66714117, 29331827,
		Inger protein PZF - mouse	Zinc finger, C2H2 type		265007, 264766, 56181562, 18108359, 18108365, 18108370, 18108381
099	81897922 (1319, 1320)			UNCLASSIFIED	264757
199	80026023 (1321, 1322)	80026023 (1321, 1322) Novel Protein sim. GBank	Contains protein domain (PF00874) - UNCLASSIFIED	UNCLASSIFIED	264510, 265009, 264600, 264602, 264603.
		gi 134180 sp P15401 SACY_BACSU - LEVANSUCRASE	Transcriptional antiterminator bglG		264604, 264605, 32833986, 18108376,
		AND SUCRASE SYNTHESIS OPERON ANTITEDMINATOR	family		264636, 18108387, 22279000
Ş	1004 00004 40000				
700	20463/31 (1323, 1324)	zu4b3731 (1323, 1324) Novel Protein sim. GBank gil4545229 gb AAD22450.1JAF11618 - (AF116183) SecA		UNCLASSIFIED	264605
		homolog [Actinobacillus actinomycetemcomitans]			
663	20628080 (1325, 1326)	Novel Protein sim. GBank gil5689250 dbj BAA82881.1 -		dehydrogenase	264605
		(AB024335) similar to orf5 [Comamonas testosteroni]			
664	80508512 (1327, 1328)	80508512 (1327, 1328) Novel Protein sim. GBank gi 1552848 dbj BAA17766 -		UNCLASSIFIED	264769
		(D90909) DNA photolyase [Synechocystis sp.]			
999	80079053 (1329, 1330)	80079053 (1329, 1330) Novel Protein sim. GBank		isomerase	264600
		gi 116841 sp P21640 COBJ_PSEDE - PRECORRIN-3B C17			
		METHYLTRANSFERASE (PRECORRIN-3			
		METHYLTRANSFERASE) (PRECORRIN-3 METHYLASE)			
999	79603142 (1331, 1332)	79603142 (1331, 1332) Novel Protein sim. GBank gij3261829jembjCAB10927 -		glycoprotein	264907, 265007
		(298260) hypothetical protein Rv1230c [Mycobacterium Imberculosis]			
299	94631802 (1333, 1334)	94631802 (1333, 1334) Novel Protein sim. GBank gil5688851 dbi BAA82702.1 -		UNCLASSIFIED	264689, 264602, 264593
		(AB017438) Orf5 [Streptomyces coelicolor]			
999	82051891 (1335, 1336)	82051891 (1335, 1336) Novel Protein sim. GBank gil3581853 emb CAA20809 -	Contains protein domain (PF00453) - ribosomalprot	ribosomalprot	264905, 264906, 264908, 264600, 264601,
		(AL031541) 50S ribosomal protein L20 [Streptomyces	Ribosomal protein L20		264603, 264605, 264760, 264689, 264636,
		coelicolor]			264638, 264639
699	12967154 (1337, 1338)			UNCLASSIFIED	264637
029	80238549 (1339, 1340)	80238549 (1339, 1340) Novel Protein sim. GBank gi 2582531 (AF026444) - 2-		synthase	264905, 264906, 264908, 264601, 264762,
		isopropylmalate synthase [Streptomyces coelicolor]			264766, 264689, 264638, 18108385, 264486
671	79601368 (1341, 1342)		Contains protein domain (PF00023) - UNCLASSIFIED Ank repeat	UNCLASSIFIED	264690, 264692, 264693, 264636, 18108387
672	79834371 (1343, 1344)	79834371 (1343, 1344) Novel Protein sim. GBank gi 2114430 (U92703) - Olf-1/EBF-		transcriptfactor	264910, 265017
		like-3 transcription factor [Mus musculus]			
673	82285798 (1345, 1346)	82285798 (1345, 1346) Novel Protein sim. GBank			264759
		gi 4589285 gb AAD26430.1 AF13515 - (AF135154) ferric alcaligin siderophore receptor [Bordetella pertussis]			
674	79199259 (1347, 1348)			UNCLASSIFIED	264629

675	87895870 (1349, 1350)	87895870 (1349, 1350) Novel Protein sim. GBank	Contains protein domain (PF01820) - UNCLASSIFIED		264488, 22278999, 66714117, 264508,
		gi 4980755 gb AAD35347.1 AE00170 - (AE001708) D-	D-ala D-ala ligase		.264511, 265008, 60433438, 264600, 264601, 1 264601, 264601, 264601, 264602, 264763
		alanındO-alanınd ilgase (Thermologa manınd)			264687, 264769, 60431602, 18108374,
878	78800607 (4264 4360)	Manual Brother of the Charle			204030, 204038
<u> </u>	(1001, 1005)	NOVER FIOLERI SIIII. GBAIRI gi 1723566 sp Q10479 YDF7_SCHPO - PUTATIVE GLUCOSYLTRANSFERASE C17C9 07			010007
229	21644312 (1353, 1354) Novel Protein sim. GF	Novel Protein sim. GBank gi[687208 (U03976) - dynein heavy chain isolyne 5C (Tripneustes gratilla)		ATPase_associated 264591, 264632	264591, 264632
829	84225200 (1355, 1356)	84225200 (1355, 1356) Novel Protein sim. GBank gil1586274 prfl 2203365A -	Contains protein domain (PF00053) - laminin	laminin	264758, 264682, 264557
		laminin alpha5 [Mus musculus]	Laminin EGF-like (Domains III and		
679	79868855 (1357, 1358) Novel Protein sim. Gl	Novel Protein sim. GBank gil3928723lemblCAA22219l -		UNCLASSIFIED	22278996, 264693
		(AL034355) putative ABC transporter [Streptomyces			
		coelicolor]			
989	20726424 (1359, 1360)				264600, 264602
681	94322017 (1361, 1362) Novel Protein sim. Gl	Novel Protein sim. GBank	Contains protein domain (PF00053) - Iaminin	laminin	264102, 264907, 264908, 265006, 264693,
		gi 5174493 ref NP_006050.1 pLAMC - Iaminin, gamma 3	Laminin EGF-like (Domains III and V)		263972, 83373044, 264566
68	11302478 /1383 1384)			Calabo V ON I	264606
7 6	11332470 (1303, 1304)			3	204090
983	80083680 (1365, 1366)	Novel Protein sim. GBank	Contains protein domain (PF00782) - phosphatase		264634
		gil4758208 ref NP_004081.1 pDUSP - dual specificity	Dual specificity phosphatase,		
		phosphatase 3 (vaccinia virus phosphatase VH1-related)	catalytic domain		
684	20465367 (1367, 1368)	20465367 (1367, 1368) Novel Protein sim. GBank gij5420387 emb CAB45679.1 - (AJ243459) proteophosphoglycan [Leishmania major]			264605
685	80246735 (1369, 1370)	_			264909, 263967, 263981
989	79208606 (1371, 1372)				264631
687	80085629 (1373, 1374)				264693, 264635
688	79853412 (1375, 1376) Novel Protein sim G	Novel Protein sim GBank gil 2688962 (AF027768) - 1 spA		pentidase	264907 264638
3	(010) (010)	Serratia marcescens		populasa	20101; 201000
689	88064256 (1377, 1378)	88064256 (1377, 1378) Novel Protein sim. GBank gij3046931 (AF049330) - PPAR	Contains protein domain (PF00076) -	UNCLASSIFIED	264906, 264907, 265007, 265009, 60433438,
		gamma coactivator [Mus musculus]	RNA recognition motif. (a.k.a. RRM,		21906754, 264760, 18108358, 21906766,
			RBD, or RNP domain)		21906769, 265021, 18108361, 263974,
					18108379, 264557, 18108385, 22279002
069	80389750 (1379, 1380)	80389750 (1379, 1380) Novel Protein sim. GBank gi[2498941[sp]Q15428]SP62_HUMAN - SPLICEOSOME		UNCLASSIFIED	264510, 264511, 264764, 264769
100	04054200 44304 43001				001101
8 8	01034392 (1301, 1302)			1	70473/
285	(83608936 (1383, 1384) Novel Protein sim. Gi (AJ243459) proteoph	Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]	Contains protein domain (PF00097) - UNCLASSIFIED Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	55812038, 55811957, 265018, 55811150, 18108351, 264908, 60431528, 264594
693	79586116 (1385, 1386)	79586116 (1385, 1386) Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 Human herpesvirus 6		UNCLASSIFIED	264635
694	82455983 (1387, 1388)	82455983 (1387, 1388) Novel Protein sim. GBank			22278996, 264510, 264602, 264603, 264762,
		gij267327 sp Q01033 VG48_HSVSA - HYPOTHETICAL GENE 48 PROTEIN			264687, 264769, 264693

264907	264692	264636	18108392, 22278994, 22278998, 265008, 265018, 264681, 18108354, 264684, 264685, 264686, 264687, 264689, 21906769, 18108361, 264691, 264692, 55810764, 264635, 18108381, 18108382, 83373044, 18108388	265011	264629	264910	264691	264909	264905, 264907	264636	264634 264567	254400	704430	264564	264605	264591, 264594, 264595	264604	264604
UNCLASSIFIED		dehydrogenase	kinase	UNCLASSIFIED	UNCLASSIFIED		UNCLASSIFIED	protease	UNCLASSIFIED	UNCLASSIFIED		CHICION	UNCLASSIFIED	isomerase	UNCLASSIFIED	struct	UNCLASSIFIED	ubiquitin
			Contains protein domain (PF00093) - von Willebrand factor type C domain					Contains protein domain (PF00353) - protease Hemolysin-type calcium-binding proteins										
79639423 (1429, 1430) Novel Protein sim. GBank gil1789035 (AE000352) - orf. Inypothetical protein [Escherichia coli]	1432)	79491842 (1433, 1434) Novel Protein sim. GBank gijz494074 sp P55653 GABD_RHISN - PROBABLE SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+) (SSDH)	94319658 (1435, 1436) Novel Protein sim. GBank gi]3873679Jemb CAA94886 - (271178) similar to pro-collagen domains; cDNA EST EMBL:D27978 comes from this gene; cDNA EST EMBL:D27977 comes from this gene; cDNA EST EMBL:D34199 comes from this gene; cDNA EST EMBL:D34199 comes from this gene; cDNA EST EMBL.	,		444) Novel Protein sim. GBank gi[498253 (U02372) - integrase [Vibrio cholerae]	19755599 (1445, 1446) Novel Protein sim. GBank gi[2253054 emb CAB10705 - (297559) hypothetical protein Rv2114 [Mycobacterium tuberculosis]	10126494 (1447, 1448) Novel Protein sim. GBank gil4063015 (AF083061) - protease PrtA [Pseudomonas fluorescens]	1450)	1452)	13522872 (1453, 1454) 20268471 (1455, 1456) Novel Protein sim. GBank gil2633910lemblCAB13411I -	(299112) similar to hypothetical proteins [Bacillus subtilis]	1400)	19900373 (1459, 1460) Novel Protein sim. GBank gi 2494660 sp Q45291 GALE_BRELA - UDP-GLUCOSE 4- EPIMERASE (GALACTOWALDENASE) (UDP- GALACTOSE 4-EPIMERASE)	1462) Novel Protein sim. GBank gi 1146192 (L47838) - putative [Bacillus subtilis]		1466)	20435987 (1467, 1468) Novel Protein sim. GBank gij3184080jembjCAA19336j - (AL023781) hypothetical protein [Schizosaccharomyces [pombe]
79639423 (1429, 14	79559072 (1431, 1432)	79491842 (1433, 14	94319658 (1435, 14	17679564 (1437, 1438) 79841684 (1439, 1440)	15020180 (1441, 14	9862603 (1443, 1444)	19755599 (1445, 14	10126494 (1447, 14	79878679 (1449, 1450)	13086282 (1451, 1452)	13522872 (1453, 1- 20268471 (1455, 14	11203753 (1457 1458)	1,293733 (1437, 14	19900373 (1459, 1,	80058750 (1461, 1	80258175 (1463, 14	20446839 (1465, 1466)	20435987 (1467, 1-
715	716	717	718	719	721	722	723	724	725	728	727	720	ş	730	731	732	733	734

264594	264636		265008	264910	264907, 264764, 264634, 264637					18108397, 264511, 264690, 264628, 264638, 264692, 264639, 264766	264906	264600	66714117, 264905, 264509, 264906, 264907, 264907, 264908, 264909, 264511, 264910, 265011, 264981, 264281, 264687, 264687, 264768, 264769, 21906768, 35695917, 264691, 264693, 264628, 264634, 264635, 264639, 26182323, 83373044	264690
	transport	UNCLASSIFIED		oxidase		UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIE			glycoprotein	synthase	oxidase	
	Contains protein domain (PF00528) - transport Binding-protein-dependent transport systems inner membrane component	Contains protein domain (PF00615) - UNCLASSIFIED Regulator of G protein signaling domain				Contains protein domain (PF01883) - UNCLASSIFIED Domain of unknown function		Contains protein domain (PF00145) - UNCLASSIFIED	C-5 cytosine-specific DNA methylase				Contains protein domain (PF00018) - oxidase SH3 domain	
11607959 (1469, 1470) Novel Protein sim. GBank gild01582[splP27432]YICE_ECOLI - HYPOTHETICAL 48.9	_			79833670 (1477, 1478) Novel Protein sim. GBank gi[2506667]spp33225[TORA_ECOLI - TRIMETHYLAMINE- N-VOXIDE REDUCTASE PRECURSOR (TMAO REDICTASE) (TRIMETHYLAMINE OXIDASE)		78627273 (1481, 1482) Novet Protein sim. GBank gij3261828 emb CAB10925 - (288260) mm [Mycobacterium tuberculosis]	82393795 (1483, 1484) Novel Protein sim. GBank gij3877494 emb CAA88472.1 - [Z48583] ATP binding protein with similarity to the CDC48/PAS1/SEC18 family, cDNA EST EMBL:D65037 comes from this gene; cDNA EST EMBL:D68340 comes from this gene; cDNA EST EMBL:D65048 comes from this gene; cDNA EST EMBL:D65048 comes from this	gene; CDNA ESI EMBL. Docto	METHYLTRANSFERASE BANI) (M.BANI)		Novei Protein sim. GBank gil78921 pir 504846 - UDP-N- acetylmuramoylalanyl-D-glutamyl-2, 6-diaminopimelateD- alanyl-D-alanine ligase (EC 6.3.2.15) precursor - Fscherichia coli	Novel Protein sim. GBank gil 3386354 (AF074705) - pyochelin synthetase IPseudomonas aeruginosal	94322044 (1493, 1494) Novel Protein sim. GBank gil 2887411 dbi BAA24848 - (AB007878) KIAA0418 [Homo sapiens]	
607959 (1469, 1470)	1471, 1472)	78945340 (1473, 1474)	17895353 (1475, 1476)	9833670 (1477, 1478)	19881557 (1479 1480)	9827273 (1481, 1482)	2393795 (1483, 1484)	7300051 11485 1486)	(no. 1 - 701) 1000000	80230421 (1487, 1488)	9841963 (1489, 1490)	11073229 (1491, 1492) Novel Protein sim. pvochelin svntheta:	из22044 (1493, 1494)	1000 1 300 10 0000 1001
735 11	736 10	737 78	╅	739 76	740		742 82	743		744 8	745 9	746 1	747	

749					,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	Г
750	20296427 (1499, 1500)	Novel Protein sim. GBank		UNCLASSIFIED	264604	- 1
		qi 1169727 sp P44			264600	
		FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (FAPY.				
751	21636169 (1501, 1502)	21636169 (1501, 1502) Novel Protein sim GBank				
		gl53600681gblACA22851.1(AF15968 - (AF159689) Sefinethreonine kinase PKN3 (Myxococcus xanthus)	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	kinase	264605, 264559	
752	82450366 (1503, 1504)	82450366 (1503, 1504) Novel Protein sim. GBank	Coccool of the city of the contract			
		gil 1686621819 2011. gil 1686621819 2011. ADENOSYLMETHIONINE-B-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)	Contains protein domain (P+0202) - Aminotransferases class-til pyridoxal phosphate	UNCLASSIFIED	264508, 264907, 264510, 265011, 264762, 284689, 35695855, 264638, 18108387	
753	80508718 (1505 1506)	80508718 (1505 1506) Novel Protein sim CBank				
		NOVEL TUGEIL SIII. GBAIK 91/2851530[sp]932399[YHGE_BACSU - HYPOTHETICAL 84.1 KD PROTEIN IN HEMY-GLTT INTERGENIC REGION (ORFB)		UNCLASSIFIED	264909, 264600, 264602, 264604, 264760, 264769, 264634	
754	95083741 (1507, 1508)	-		CLILLIOO VIOLATI		—,
				UNCLASSIFIED	264508, 264906, 264907, 264908, 264909, 264759, 264602, 264764, 264769, 264628	
					264629 264630 264632 264634 264635	
7,66	00405440 (4500 4540)				264637, 264638, 83373044, 18108385	
	60162449 (1509, 1510)	0103, 1310)		UNCLASSIFIED	264448 264690	_
8	94631686 (1511, 1512)	Novel Protein sim. GBank gil3449276 emblCAA204201 -		Ī	264760 264690 264620	_
		(AL031317) putative dehydrogenase (Streptomyces			204705, 204065, 204036, 204039	
757	79468533 (1513 1514)	COERICOIOL				
75.0	78082478 (4848 4648)	0		UNCLASSIFIED	264682, 264685	_
3	(1313, 1318)	(AJ001201) Novel Protein Sim. GBank gild580331 emb CAB40107.1 - (AJ001206) putative gilycogen debranching enzyme [{Streptomyces coelicolor]		amylase	265007, 18108387, 265007, 18108387	
759	79475667 (1517, 1518)	Novel Protein sim. GBank gi 2911858 (AF047659) - No		UNCLASSIFIED	264684 264686	
25	07270000 /1540 4500	definition line found [Caenorhabditis elegans]		_		
3	0/02000 (1019, 1020)	0/020000 (1319, 1320) Novel Protein sim. GBank gij3451312 emb CAA20449 .	Contains protein domain (PF00122) - transport		29331822, 264908, 52644045, 56182435,	
			E1-E2 ATPase		60170831, 21906754, 265017, 265019, 264681, 264687, 264688, 21006766	
					201001; E01001; E01001, E1800100;	
					21906768, 265020, 265021, 265022, 264635,	
761	79877966 (1521, 1522)			A INC. A COLLING	2000	_
762	80023563 (1523, 1524)	80023563 (1523, 1524) Novel Protein sim. GBank gij3327158jdbjjBAA31647i -		T	264007 264503 205220	_
163	2027	(AB014572) KIAA0672 protein [Homo sapiens]			204307, 204383, 203020	_
ê	zuz94813 (1525, 1526)	20294813 (1525, 1526) Novel Protein sim. GBank			264600	
		gil4981266 gb AAD35822.1 AE00174 - (AE001744)				
		impoporysacciarine core biosynmesis protein KdtB [Thermotoga maritima]		•		
764	39515024 (1527, 1528)				264603	
					20010	

765	80025347 (1529, 1530)	80025347 (1529, 1530) [Novel Protein sim. GBank gil3845093 (AE001371) -		struct	264905 284906 264594 264686 33657033
		erythrocyte membrane protein PfEMP3 [Plasmodium			100000000000000000000000000000000000000
		falciparum			
99/	82417404 (1531, 1532)				264605, 264762, 18108374
1 92	10296742 (1533, 1534) Novel Protein sim. GI hypothetical protein o	Novel Protein sim. GBank gi 541121 pir S40827 - hypothetical protein o300 - Escherichia coli		UNCLASSIFIED	264907
292	79416080 (1535, 1536)			UNCLASSIFIED	264592, 264595
769	80086554 (1537, 1538)	80086554 (1537, 1538) Novel Protein sim. GBank gij2982501 emb CAA06164 - (AJ004832) neuropathy tardet esterase IHomo saniensi		esterase	55810764, 264559
770	80417847 (1539, 1540)	80417847 (1539, 1540) Novel Protein sim. GBank gi[283437 pir [S27850 -		UNCLASSIFIED	264905 264907 264828 264909 265010
		hypothetical protein - Trypanosoma cruzi (fragment)			264766, 264628, 264629, 264634, 264636, 264555
771	95329509 (1541, 1542)	95329509 (1541, 1542) Novel Protein sim. GBank	Contains protein domain (PF00097) - UNCLASSIFIED	UNCLASSIFIED	56182575, 35696286, 264259, 66714117.
		gil4769004[gb]AAD29715.1 AF14059 - (AF140598) ring-box Zinc finger, C3HC4 type (RING	Zinc finger, C3HC4 type (RING		264107, 66712502, 56182435, 264112,
		protein 1 [Homo sapiens]	finger)		55812038, 87168559, 264288, 21906766, 33657023, 65274620, 65274791, 18108381
772	78971362 (1543, 1544)			UNCLASSIFIED	264910
773	78945363 (1545, 1546)			UNCLASSIFIED	265020
774	79856129 (1547, 1548) Novel Protein sim. GE	Novel Protein sim. GBank gi[5531324 emb]CAB51045.1 -		UNCLASSIFIED	264909
		(AJ009579) putative alkane 1-monooxygenase [Pseudomonas fluorescens]			
775	20620141 (1549, 1550)			INCIACCIETED	25.155
776	78942693 (1861 1862)			CINCLASSII ICD	20010
	(301, 101)		Contains protein domain (PF 01005) - Hepatitis C virus non-structural protein NS4a		265019
111	(79960378 (1553, 1554)	79960378 (1553, 1554) Novel Protein sim. GBank	Contains protein domain (PF01344) - protease	profease	21906754, 265020, 60170615, 264691
		gil4505461 [ref[NP_003624.1]pNRPB - nuclear restricted protein, BTB domain-like (brain)	Kelch motif		
778				UNCLASSIFIED	264511
779	80054024 (1557, 1558)				264603
780	95288987 (1559, 1560)	95288987 (1559, 1560) Novel Protein sim. GBank gil1144520 (U34956) -		synthase	264907 264600 264601 264602 264603
		phosphoribosylformylglycinamidine synthase [Mycobacterium tuberculosis]			264604, 264605, 264486
781	80250049 (1561, 1562)				264905, 264907, 265010, 264600, 264601, 18108362, 18108374, 264556
782	8758529 (1563, 1564)	Novel Protein sim. GBank gil4155447 (AE001517) - proline/hefaine transander Helicoharte pulori 1001		UNCLASSIFIED	264605
783	16410791 (1565, 1566)	foo wild proposed by the second			26.60.20
784	80051197 (1567, 1568)			INCI ASSIFIED	203020 264635 33657023 20331828 265017
į					264565, 264568
8	56073541 (1569, 1570)	560/3541 (1569, 1570) Novel Protein sim. GBank gij3451335 (AC005525) - F22162_1 [Homo sapiens]	Contains protein domain (PF00047) - struct Immunoglobulin domain	struct	35696052, 264604
786	20438842 (1571, 1572)	20438842 (1571, 1572) Novel Protein sim. GBank		transport	264603
		gil136748 sp P10905 UGPA_ECOLI - SN-GLYCEROL-3- PHOSPHATE TRANSPORT SYSTEM PERMEASE			
787	00058054 (1579 1574)	PROTEIN UGPA			
è	00200204 (1013, 1014)			UNCLASSIFIED	264593

								_	_	_		_			_		_		
264909, 264602, 264603, 264769, 264638	265007	22278999, 29331824, 264828, 87168559, 265018, 21906765, 21906767, 21906768, 21906769, 265020, 264692, 22279000, 264563	264601, 264605, 264636	265021, 264631, 264635, 264556	264693	264603, 264604, 264910, 264762, 264906, 264639, 264909, 264757	264488, 35696052, 264905, 264907, 265010, 35696423, 264636	60432229, 55811150, 264630, 264637, 264565	22278998, 264093, 264094, 66714117, 21906767, 21906769, 265020, 265022	264687, 264768, 264693	264909, 264910, 264636, 264638	264488	264508, 264511, 265006, 265009, 264769, 264567, 264486	264905, 264693	264605	264510	264508	264769	264764
synthase	UNCLASSIFIED	glycoprotein		struct	homeobox	UNCLASSIFIED	carboxylase		struct	transcriptfactor	hydrolase	UNCLASSIFIED	transport		struct	kinase		Iransferase	
		Contains protein domain (PF00047) - glycoprotein Immunoglobulin domain			Contains protein domain (PF00046) - homeobox Homeobox domain													Contains protein domain (PF00185) - Iransferase Aspartate/ornithine carbamoyltransferase	
80507844 (1575, 1576) Novel Protein sim. GBank gil2746079 (AF015310) - BTH1 [Brassica napus]	j		94651627 (1581, 1562) Novel Protein sim. GBank gi 5689948 emb CAB51985.1 - (AL109663) putative isoleucyl-tRNA synthetase Streptomyces coelicolor A3(2)]	u G	79638730 (1585, 1586) Novel Protein sim. GBank gi 1345408 dbj BAA05046 - (D26046) AT motif-binding factor (Mus musculus)	81839294 (1587, 1588) Novel Protein sim. GBank gi[105884 pir S24023 - dopamine receptor D4 - human (fragment)	80074988 (1589, 1590) Novel Protein sim. GBank giļ1877334 emb CAB07082 - (Z92771) birA (Mycobacterium tuberculosis)	_	87771781 (1593, 1594) Novel Protein sim. GBank gi[2995447 emb CAA71519 - (Y10495) CDV-1R protein [Mus musculus]	95, 1596)	79557816 (1597, 1598) Novel Protein sim. GBank gi[4467250]emb CAB37575] - (AL035569) probable Glu-tRNA Gln amidotransferase subunit [Streptomyces coelicolor]	99, 1600)	80499399 (1601, 1602) Novel Protein sim. GBank gi[2791517]emb CAA16054] - (AL021246) hypothetical protein Rv2477c [Mycobacterium [tuberculosis]	79834598 (1603, 1604) Novel Protein sim. GBank gi 4887211[gb AAD32237.1 AF14744 - (AF147449) peniciliin binding protein 1B Pseudomonas aeruginosa]		10174239 (1607, 1608) Novel Protein sim. GBank gij1176152[sp P44607[YHAD_HAEIN - HYPOTHETICAL PROTEIN HI0091		80484113 (1611, 1612) Novel Protein sim. GBank gi[2764612 emb CAA04683] - (AJ001330) omithine transcarbamoylase [Lactobacillus sakei]	80381812 (1613, 1614) Novel Protein sim. GBank gil2833311 sp Q21828 YNFD_CAEEL - HYPOTHETICAL 18.9 KD PROTEIN R07E5.13 IN CHROMOSOME III
80507844 (157	17294715 (157	86284406 (157	94651627 (158	80058786 (158	79638730 (158	81839294 (158	80074988 (158	86669451 (1591, 1592)	8777781 (159	79865209 (159	79557816 (155	(79970189 (1599, 1600)	80499399 (160	79834598 (160	20467520 (1605, 1606)	10174239 (160	79599993 (1609, 1610)	80484113 (161	80381812 (161
788	789	062	791	792	793	794	795	962	797	798	662	800	801	802	803	8 4	805	908	807

808	35106817 (1615, 1616) Novel Protein sim G	Novel Protein sim GBank		100000	
		gigg13092[sp[Q46170]ARCD_CLOPE - ARGININFORMTHINE ANTIPORTER		Jodena	Z049U9, Z040UZ, Z19Ub/64, 181U83/4
808	81454254 (1617 1618)	81454254 (1617 1618) Novel Protein sim CRank	A FOOD TO THE STORE OF STORE OF		
		dilagrander Comment Co	Contains protein domain (Prouz.4) - UNCLASSIFIED	UNCLASSIFIED	264508, 264906, 264909, 265007, 264910,
		BISPHOSPHATE ALDOLASE CLASS I (FBP ALDOLASE)	Class-I		204730, 204000, 204002, 204603, 204603, 264687, 264769, 264689, 264636, 264486
810	80192761 (1619, 1620)	80192761 (1619, 1620) Novel Protein sim. GBank	Contains protein domain (PF00248) - reductase	reductase	264369
·		gij401472lsp P30863 YAFB_ECOLI - HYPOTHETICAL OXIDOREDUCTASE IN ASPU-MLTD INTERGENIC	Aldo/keto reductase family		
118	80079280 (1621 1622)	SOCIAL		This a contrar	0117
812	10297654 (1623, 1624)			UNCLASSIFIED	264603
813	79612280 (1625, 1626)			ONCOUNT OF THE PROPERTY OF THE	760407
814	80473427 (1627, 1628) Novel Protein sim Cl	Nowel Protein sim		UNCLASSIFIED	264906
	024, 1921, 1920	INOVEL Frotein sim. GBank gij 146168 (J01617) - glutaminyl- IRNA synthetase [Escherichia coli]		synthase	264905, 264602, 264605, 264682, 264687,
815	95419513 (1629, 1630)	95419513 (1629, 1630) Novel Protein sim. GBank gil4589652 dbj BAA76848.1 -		UNCLASSIFIED	264488, 22278998, 22278999, 29331822,
		(AB023221) KIAA1004 protein [Homo sapiens]			29331824, 29331825, 29331827, 29331828,
					29146499, 264905, 264908, 265007,
					33657402, 60433356, 60433438, 264758,
					265011, 265017, 265018, 265019, 264369,
					264288, 264685, 21906765, 21906767,
					265020, 265021, 264692, 65274620,
					33657109, 264629, 18108376, 264635,
970	40004040404				264638, 60170394, 56182323, 264564
0 2	19001910 (1631, 1632)				264600
5	95293316 (1633, 1634) Novel Protein sim. Gl			UNCLASSIFIED	264595
		(Z83866) hypothetical protein Rv3069 [Mycobacterium tuberculosis]			,
818	90938190 (1635, 1636) Novel Protein sim. GE	Novel Protein sim. GBank gi[1477468 (U35244) - vacuolar			65274572, 22278999, 60424269, 35696052.
		protein sorting homolog r-vps33a [Rattus norvegicus]			55812038, 21906768, 55811957, 35695917,
					33657023, 18108370, 18108374, 55810764,
810	80254977 (1637 163B)	80254977 (1637 1638) Navel Drateis sim CBank million 25014 hills 4 108201			33080423, 33811370, 204030
	(201, 1001)	(D64006) ABC transporter [Synechocystis sp.]		transport	264565
820	80059688 (1639, 1640)	80059688 (1639, 1640) Novel Protein sim. GBank		UNCLASSIFIED	264600, 264602, 264604
		gi 586814 sp P37484 YYBT_BACSU - HYPOTHETICAL 74.3 KD PROTEIN IN RPI I-COTF INTERGENIC REGION			
821	79762590 (1641, 1642)			UNCLASSIFIED	264910
822	80215310 (1643, 1644)			Π	264510 264594 264637
823	94992299 (1645, 1646) Novel Protein sim. GE	Novel Protein sim. GBank gij3878400 emb CAA95828 -			264509, 264687, 264691
		(Z71264) predicted using Genefinder; Weak similarity to			
		Mouse T-complex-associated-testes-expressed-1 protein			
		from this gene; cDNA EST EMBL:D33617 comes from this gene; cDNA EST			
824	80411171 (1647, 1648) Novel Protein sim. GE	Novel Protein sim. GBank gil1370076 emb CAA66887 -	Contains protein domain (PF01429) -		264910, 264763, 264769, 264693
		(Asocos) type i prosopnila meianogaster	Methyl-CpC binding domain		

UNCLASSIFIED 264592	264605	264603	UNCLASSIFIED 52645080, 264908, 264909, 264509, 264906, 264909, 264909, 264910, 264901, 264909, 264909, 264910, 264591, 264758, 264600, 264766, 264768, 21906768, 35895917, 264631, 264628, 264630, 264631, 264632, 264636, 264639, 264637, 264535, 264598, 264593, 264583, 264586, 264686, 264	UNCLASSIFIED 264758, 264634	UNCLASSIFIED 264603	UNCLASSIFIED 66714117, 264910, 264639	UNCLASSIFIED 264906	UNCLASSIFIED 264602	4) - kinase 264600, 264602, 264769, 264689, 264636	264768, 263994, 21906767, 264910, 264632, 264635, 264259, 264639, 264693, 83373044, 264758, 35696052, 22279002, 264508, 264905, 264906, 264448, 263972, 264908, 264909	Iransport 35695917, 264557	UNCLASSIFIED 264595	UNCLASSIFIED 264687	UNCLASSIFIED 264906	7) - interleukinrecept 264762, 264556
									Contains protein domain (PF00224) - kinase Pyruvate kinase						Contains protein domain (PF00097) - interleukinrecept Zinc finger, C3HC4 type (RING
20638600 (1649, 1650) Novel Protein sim. GBank gij3025132 sp P77391 YEAG_ECOLI - HYPOTHETICAL 14.5 KD PROTEIN IN GAPA-RND INTERGENIC REGION	651, 1652) Novel Protein sim. GBank gij3242281 emb CAA16669] - (AL021646) hypothetical protein Rv3202c [Mycobacterium tuberculosis]	80054207 (1653, 1654) Novel Protein sim. GBank gij3417424jembjCAA20312j - (AL031261) putative transport protein [Schizosaccharomyces pombe]	95106322 (1655, 1656) Novel Protein sim. GBank gil433692 gb AAD17897 - (AF101361) Abnormal X segregation [Drosophila melanogaster]	657, 1658)	20396091 (1659, 1660) Novel Protein sim. GBank gij3820539 (AF080002) - UDP-N-acetylmuramyl tripeptide synthetase MurC [Heliobacillus mobilis]	561, 1662)	19536322 (1663, 1664) Novel Protein sim. GBank gi 1870004 emb CAB06855 - (292639) hypothetical protein Rv1024 Mycobacterium tuberculosis]	20726654 (1665, 1666) Novel Protein sim. GBank gi[2500056 sp Q46267 PFLA_CLOPA - PYRUVATE FORMATE-LYASE ACTIVATING ENZYME		669, 1670 <u>)</u>	66126552 (1671, 1672) Novel Protein sim. GBank gil699315 (U15184) - phosphate transport protein PSTC [Mycobacterium leprae]	673, 1674)	79184203 (1675, 1676) Novel Protein sim. GBank gij728867jspIP40602/APG_ARATH - ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR	677, 1678) Novel Protein sim. GBank gi[2496533 sp Q50598 Y0D8_MYCTU - HYPOTHETICAL 69.9 KD PROTEIN CY1A11.08	80059851 (1679, 1680) Novel Protein sim. GBank gij4557753 ref NP_000372.1 pMID1 - midline 1 protein
825 20638600 (16	826 11075047 (1651, 1652)		828 95106322 (16	829 81742215 (1657, 1658)		831 87112435 (16						837 79450450 (1673, 1674)			840 80059851 (16

2	803/6318 (1681, 1682) Novel Protein sim. GB. gil139805[sp P08045 x	Novel Protein sim. GBank gi 139805 sp P08045 XFIN_XENLA - XFIN PROTEIN	Contains protein domain (PF00096) - transcriptfactor Zinc finger, C2H2 type	transcriptfactor	264764
842	80078724 (1683, 1684) Novel Protein sim. GBa	Novel Protein sim. GBank gi 2114321 db BAA20037 -	Contains protein domain (PF00569) - UNCLASSIFIED	UNCLASSIFIED	264905, 264908, 265008, 265009, 18108374,
		(D88733) membrane glycoprotein [Equine herpesvirus 1]	Zinc finger present in dystrophin, CBP/p300		56182323, 264558
843	87002847 (1685, 1686)	87002847 (1685, 1686) Novel Protein sim. GBank gij3882325/dbjjBAA34522.1 - (AB018345) KIAA0802 protein [Homo saplens]	Contains protein domain (PF00170) - struct bZIP transcription factor	struct	264091, 29331825, 264906, 264768, 264563
844	17941439 (1687, 1688)	17941439 (1687, 1688) Novel Protein sim. GBank gi[2224721 dbj BAA20844 - [AB002388) KIAA0390 [Homo sapiens]	Contains protein domain (PF00096) - transcriptfactor Zinc finger, C2H2 type	transcriptfactor	265011
842	18346844 (1689, 1690)	+-			264629
848	79863441 (1691, 1692)	79863441 (1691, 1692) Novet Protein sim. GBank gi[625679 pir A36929 - virulence regulatory protein VsrB - Pseudomonas solanacearum		kinase	264907
847	78695348 (1693, 1694)				264909
848	78489365 (1695, 1696)			UNCLASSIFIED	265020
849	79756367 (1697, 1698)			UNCLASSIFIED	264566
820	79817849 (1699, 1700)		Contains protein domain (PF00120) - UNCLASSIFIED	UNCLASSIFIED	264909
		gij3183245 sp P78061 YCJK_ECOLI - PUTATIVE GLUTAMINE SYNTHETASE (GLUTAMATE-AMMONIA LIGASE)	Glutamine synthetase		
851	95320333 (1701, 1702)	95320333 (1701, 1702) Novel Protein sim. GBank	Contains protein domain (PF01608) -		264488, 52644507, 264489, 18108398,
		gi[5454130 ref NP_006280.1 pTLN - talin	I/LWEQ domain		65274572, 56182575, 22278994, 22278995.
					22278996, 35696286, 22278997, 22278998,
					22278999, 20281171, 264490, 264259,
					52645080, 29331822, 29331824, 66714117,
					29331825, 60432289, 29331826, 29331827,
					35696052, 29331828, 29146498, 29146499,
					264107, 264905, 264906, 264907, 264908,
	-				52644045, 56182435, 265006, 265007,
					265008, 265009, 264910, 60432229.
					60431735, 60433356, 33657402, 60433438.
					264595, 264758, 264759, 21906754,
					33109954, 52644296, 265010, 265011,
_					87168559, 265017, 265018, 265019, 264760,
					264761, 264762, 264681, 18108351, 264763,
					264448, 264682, 264764, 264683, 18108354,
					264288, 264369, 264685, 264766, 264687.
					264768, 264769, 21906765, 21906766,
					21906767, 21906768, 29148627, 21906769,
					29148629, 55811957, 35695917, 265020,
					265021, 265022, 60170615, 52644150,
					264691, 264692, 33657023, 264693, 263966,
					33657109, 27486261, 27486262, 27486264,
					27486265, 35695763, 60431602, 18108370,
					20281069, 264629, 18108374, 18108376,
					55811576, 35696423, 35695855, 264634,
9					264635, 264636, 264555, 60431850, 264556,
700	10147300 (1703, 1704)				264691

650	12022 11705				
854	80052438 (1707 1708	20022301 (1702, 1700) 80052438 (1707, 1708) Novel Drotein vim GBank ail34038261ambic 8 8 260831			264636
	2011 (2011) 2012	(Y16136) 2-enoate reductase [Moorella thermoacetica]		reductase	264568
855	79641130 (1709, 1710)				264692
826	11594236 (1711, 1712)	16		UNCLASSIFIED	264591
857	79210165 (1713, 1714)	(1)		UNCLASSIFIED	264630, 264634
828	80248910 (1715, 1716)	((265008, 265009, 264601, 264602, 264603, 18108351
859	20296634 (1717, 1718)	(1)			264559
860	80041749 (1719, 1720)			UNCLASSIFIED	264489
861	65857045 (1721, 1722)			UNCLASSIFIED	33657023 264630
862	80079467 (1723, 1724)				264600
863	80579931 (1725, 1726	80579931 (1725, 1726) Novel Protein sim. GBank gi[2246532 (U93872) - ORF 73,		UNCLASSIFIED	264488, 18108398, 35696286, 264259,
		contains large complex repeat CR 73 [Kaposi's sarcoma-			18108351, 264288, 265021
		associated herpesvirus]			
864	94939904 (1727, 1728)	(1		UNCLASSIFIED	264259, 264112, 263974
865	80045310 (1729, 1730		Contains protein domain (PF01479) -		264635, 264600, 264636, 264591, 264602,
		(AL109732) hypothetical protein [Streptomyces coelicolor A3(2)]	S4 domain		264693
998	80162031 (1731, 1732	80162031 (1731, 1732) Novel Protein sim. GBank		transport	264288, 264557, 264558
		gil4557876 ref NP_000341.1 pABCR - ATP binding cassette			
1		-			
ğ	80062402 (1733, 1734)				264605
8	10075364 (1735, 1736)			UNCLASSIFIED	264909
869	80062406 (1737, 1738	(1)			264605, 264687, 18108374
870	80249651 (1739, 1740	80249651 (1739, 1740) Novel Protein sim. GBank gij628660 pir S37755 - Adenylyl- Iraneferace - Escherichia cyli		transferase	264601, 264636
074	002000 11211				
5	203/8295 (1/41, 1/42	20378293 (1741, 1742) Novel Protein sim. GBank gi 1708180 sp Q10602 HEMK_MYCTU - HEMK PROTEIN HOMOLOG		UNCLASSIFIED	264603
872	95197114 (1743, 1744	95197114 (1743, 1744) Novel Protein sim. GBank gi 1545959 emb CAA67763 -		UNCLASSIFIED	35696286, 22278998, 264259, 29331822,
		(X99384) paladin [Mus musculus]			29331824, 29331825, 60432289, 29331826,
					28331627, 28331626, 33086032, 264508, 264606, 264606, 264607, 264608
					204909, 204900, 204901, 204900, 204909, 1
					204310, 203000, 204311, 204312, 2030U/,
					201000, 201003, 2048 0, 20439 , 00455550, 00456474
					204330, 320403 11, 07 100474, 203010, 264603 264603 266013 266018 264606
					2040UZ, 2040U3, 203UIT, 203UIB, 2040U3, 1810R3£1 2647E4 2647E9
					10100301, 204/04, 204/00, 204/00, 62644300, 264360, 24006466, 266034
					32044229, 204709, 21900705, 205021, 364634 364604 53646430 364630 364630
					204334, 204031, 32043128, 204026, 204028, 7
					35595423, 55274791, 254531, 254532,
					204035, 204036, 204556, 204037, 204038,
					204039, 00432113, 22279000, 22279002, 264564
873	20189728 (1745, 1746	- putative	Contains protein domain (PF00096) -	UNCLASSIFIED	264595
		Outer membrane protein [Helicobacter pylori J99]	Zinc finger, C2H2 type		

891	13516879 (1781, 1782)	13516879 (1781, 1782) Novel Protein sim. GBank gil4959396[gb]AAD34331.1IAF11248 - (AF112481) RAD54B		UNCLASSIFIED	264636
		protein [Homo sapiens]			
892	87634157 (1783, 1784)	87634157 (1783, 1784) Novel Protein sim. GBank gil545526 bbs 143833 - LBP- 1b=transcription factor binding to initiation site of HIV-1 {alternatively spliced} [human, Namatwa cells, Peptide, 541 aa]		transcriptfactor	22278996, 22278999, 29331828, 35696052, 264908, 265009, 265001, 264602, 265019, 264602, 265019, 264766, 21906765, 21906766, 21906769, 265020, 265021, 56526486
893	79168037 (1785, 1786)	79168037 (1785, 1786) Novel Protein sim. GBank gilz829688jsp P80608jCYSK_MAIZE - CYSTEINE SYNTHASE (O-ACETYLSERINE SULFHYDRYLASE) (O- ACETYLSERINE (THIOL)-LYASE) (CSASE)		synthase	264689, 263967
894	[11102240 (1787, 1788)				263978
895	80239868 (1789, 1790)			UNCLASSIFIED	264508, 264600, 264555, 264559
968	79747803 (1791, 1792)				264632
897	94991923 (1793, 1794)		Contains protein domain (PF00047) - UNCLASSIFIED Immunoglobutin domain	UNCLASSIFIED	264686, 29331828, 264511
868	87895109 (1795, 1796)				56182575, 60432289, 56182435, 60432229, 55811957, 22279000, 264486
899	11100463 (1797, 1798)				264601
8	80499768 (1799, 1800)	80499768 (1799, 1800) Novel Protein sim. GBank gi 1750127 (U66480) - YncC Bacillus subkils		transport	264769, 264691, 264563
901	80502410 (1801, 1802)	80502410 (1801, 1802) Novel Protein sim. GBank gij3122879!sp 007438 SYA_MYCTU - ALANYL-TRNA SYNTHETASE (ALANINETRNA LIGASE) (ALARS)	Contains protein domain (PF01411) - synthase tRNA synthetases class II (A)	synthase	264907, 264602, 264605, 264769, 35695917, 18108376, 264563
905	80503301 (1803, 1804)	80503301 (1803, 1804) Novel Protein sim. GBank gi[3355701 emb[CAA20001] - (AL031124) 3-isopropylmalate dehydratase large subunit [Streptomyces coelicolor]	Contains protein domain (PF00330) - isomerase Aconitase family (aconitate hydratase)	isomerase	264909, 265008, 264602, 264604, 264769, 264689, 264693
903	82060206 (1805, 1806)	82060206 (1805, 1806) Novel Protein sim. GBank gi 2960120 emb CAA18018.1 -	Contains protein domain (PF00370) - kinase	kinase	35696052, 264905, 264510, 264511, 264512,
		(AL022121) glpK [Mycobacterium tuberculosis]	FGGY family of carbohydrate kinases		264605, 264760, 18108351, 264762, 264687, 264768, 264769, 264768, 264769, 264688, 21906764, 35695917, 27486262, 35695855, 264634, 284636, 264486
904	20451078 (1807, 1808)	20451078 (1807, 1808) Novei Protein sim. GBank gij728887jsp P40906 ARG COCIM - ARGINASE	Contains protein domain (PF00491) - hydrolase Arginase family	hydrolase	264604
902	9398483 (1809, 1810)	Novel Protein sim. GBank gil4567200[gblAAD23616.1 AC00716 - (AC007168) hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	264909
906	80052628 (1811, 1812)	80052628 (1811, 1812) Novel Protein sim. GBank gij3738200jemb[CAA21292] - (AL031855) putative vacuolar membrane protein [Schizosaccharomyces pombe]			264595, 264605
907	87913201 (1813, 1814)			UNCLASSIFIED	60432289, 264601, 264690
908	11754482 (1815, 1816)			UNCLASSIFIED	264638
606	20727907 (1817, 1818)	20727907 (1817, 1818) Novel Protein sim. GBank gij3868940 dbj BAA34296 - (AB015054) Alg2 [Rhizomucor pusillus]		UNCLASSIFIED	264602
910	16776206 (1819, 1820)	16776206 (1819, 1820) Novel Protein sim. GBank gil4589726 dbj BAA76883.1 - [AB003137) DnaJ homolog protein [Salix gilgiana]	Contains protein domain (PF00684) - eph DnaJ central domain (4 repeats)	eph	265009
1					

911	87454340 (1821, 1822)	87454340 (1821, 1822) Novel Protein sim. GBank gil548774 sp P35685 RL7A_ORYSA - 60S RIBOSOMAL PROTEIN 17A	(E	ribosomalprot	265010, 264604, 60432113
912	20448863 (1823, 1824)	20448853 (1823, 1824) Novel Protein sim. GBank gij2314008 gb AAD07921.11- (AE000597) CDP-diglyceride hydrolase (cdh) [Helicobacter pylori 26895]	Ė	hydrolase	264559
913	20469357 (1825, 1826)		2	UNCLASSIFIED	264604
914	79183351 (1827, 1828)	79183351 (1827, 1828) Novet Protein sim. GBank gi[417657]sp]Q03604 RIR1_CAEEL - PROBABLE RIRONI ICLEOSIDE.DIPHOSPHATE PEDI ICTASE LABGE	Contains protein domain (PF00317) - Ribonucleotide reductase	eductase	264636
		CHAIN (RIBONUCLEOTIDE REDUCTASE)			
915	87606703 (1829, 1830)	87606703 (1829, 1830) Novel Protein sim. GBank gilj5689571 db BAA83069.1 - (AB029040) KIAA1117 protein [Homo sapiens]			18108398, 22278996, 66714117, 264906, 264591, 21906768, 265020, 55811576, 264638
916	79444091 (1831, 1832)	79444091 (1831, 1832) Novel Protein sim. GBank gild186110 emb CAA71790 - (Y10831) butative integrase (Raistonia eutrophal		UNCLASSIFIED	264595
917	20195985 (1833, 1834)		2	UNCLASSIFIED	264605
918	91226795 (1835, 1836)	91226795 (1835, 1836) Novel Protein sim. GBank gil1655699[emb CAA69032]	<i>S</i> 1	synthase	264259, 29331826, 264908, 265019, 264448, 265020, 264635, 83373044
919	80436785 (1837, 1838)	80436785 (1837, 1838) Novel Protein sim. GBank gils689968jembjCAB52005.11			265006, 264512, 264600, 264602, 264604,
		(one joine profits)	•		204/00, 101003/0, 204003
920	79606095 (1839, 1840)	79606095 (1839, 1840) Novel Protein sim. GBank	<u>a</u>	peptidase	264508
		gi[168448jsp]Q05813JAMP1_STRL1 - XAA-PRO AMINOPEPTIDASE I (X-PRO AMINOPEPTIDASE I)			
921	19858634 (1841, 1842)	Novel Protein sim. GBank gij3850084(emb CAA21911.1] - (AL033389) alcohol dehydrogenase (Schizosaccharomyces nombe)		UNCLASSIFIED	264600
922	78982605 (1843, 1844)	_			265019, 22279002
923	86695830 (1845, 1846)	Novel Protein sim. GB	1	tubulin	264907, 265008, 265009, 265010, 18108351,
-		gi 267079 sp P29514 TBB6_ARATH - TUBULIN BETA-6 CHAIN			264689, 265021, 18108370, 18108374, 18108385
924	21431341 (1847, 1848)				264510
925	20630332 (1849, 1850)	20630332 (1849, 1850) Novel Protein sim. GBank	9	esterase	264603
		ACTIVATING FACTOR ACETYLHYDROLASE			
		PRECURSOR (PAF ACETYLHYDROLASE) (PAF 2-			
		ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL-PASSOCIATED PHOSPHOLIPASE A2) (LDL-PLAC2) (2-ACETYL-1-	-		
		ALKYLGLYCEROPHOSPHOCHOLINE ESTERASE) (1-			
926	79397657 (1851, 1852)	79397657 (1851, 1852) Novel Protein sim. GBank gij3882325 dbj BAA34522.1 - [/AB018345] KIAA0802 protein Homo saniens		;	55811957, 263972, 264639
927	37036201 (1853, 1854)			UNCLASSIFIED	264769

928	80070610 (1855, 1856)			dehydrogenase	264605	
8	200000000000000000000000000000000000000	dehydrogenase (AA at 2538) [Calliphora vicina]				
200	20030330 (1857, 1858)			UNCLASSIFIED	264603	Т
05.8 8.30	5496348 (1859, 1860)	Novel Protein sim. GBank gij4115936jgbjAAD03446.1			264259	1
931	10245731 (1861, 1862)	10245731 (1861, 1862) Novel Protein sim. GBank gij4490609jembjCAB38642.1	Contains protein domain (PF00317) - reductase Ribonucleotide reductase	reductase	264486	1
932	80420613 (1863, 1864)	80420613 (1863, 1864) Novel Protein sim. GBank gijs459396 emb CAB50754.1 - (AL096839) putative integral membrane transport protein [Streptomyces coelicolor]		transport	264508, 264905, 264906, 264909, 264600, 264602, 264603, 264605, 264766, 264558,	1
933	94326010 (1865, 1866)	94326010 (1865, 1866) Novel Protein sim. GBank gij5689523 dbj BAA83045.1 - (AB029016) KIAA1093 protein [Homo sapiens]	Contains protein domain (PF00627) - UNCLASSIFIED UBA domain	UNCLASSIFIED	264508, 264686, 264693, 27486261, 281808370, 65274791, 284638, 264559,	
934	80039105 (1867, 1868)	80039105 (1867, 1868) Novel Protein sim. GBank gi 119111 sp P12978 EBN2_EBV - EBNA-2 NUCLEAR PROTEIN			264369, 22279002	T
935	80063162 (1869, 1870)			UNCLASSIFIED	264604, 264605, 264693, 18108370, 18108374	
936	80026632 (1871, 1872)	80026632 (1871, 1872) Novel Protein sim. GBank gi 845686 (M32103) - ORF-27 [Staphylococcus aureus]		UNCLASSIFIED	22278995, 22278996, 264602, 264687, 32833986, 18108387	T -
937	80250273 (1873, 1874)	80250273 (1873, 1874) Novel Protein sim. GBank gi 1360669 pir CGHU1V -		synthase	263978	т—
88	80026633 (1875, 1876)	80026633 (1875, 1876) Novel Protein sim. GBank gi 2226243 emb CAA74531.1 - (Y14083) hypothetical protein [Bacillus subtilis]		UNCLASSIFIED	264602	τ
838	110/1694 (1877, 1878)				264600	$\overline{}$
046 0	94144252 (1879, 1880)	94144252 (1879, 1880) Novel Protein sim. GBank gil3560166[emb CAA20678] - (AL031525) ubiquitin carboxyl-terminal hydrolase [Schizosaccharomyces pombe]	Contains protein domain (PF00442) - ubiquitin Ubiquitin carboxyl-terminal hydrolases family 2	ubiquitin	264905, 264906, 264907, 264908, 264909, 264511, 264910, 264582, 33657402, 264596, 264760, 264683, 264766, 264768, 264769, 33657023, 33657109, 284628, 264630, 264630, 264635, 264636, 264555,	T
					264637, 264556, 264638, 264639, 83373044, 18108385, 264565, 18108391	
5 3	11398414 (1881, 1882)				264593	_
3 3	19484122 (1883, 1884)			UNCLASSIFIED	264760	_
2	acuduzad (1885, 1886) Novel Protein sim. Gl apolipoprotein N-acyl aeruginosa]	Novel Protein sim. GBank giļ4033729 (AF038595) - apolipoprotein N-acyltransferase [Pseudomonas aeruginosa]		UNCLASSIFIED	264600, 264687, 264689, 264563	
44	80216096 (1887, 1888) Novel Protein sim. GF gi 2494764 sp Q5072 GLUTAMINE-HYDRI AMIDOTRANSFERA	Novel Protein sim. GBank gi[2494764 sp Q50729 GUAA_MYCTU - GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE)		synthase	264511, 264603	
945 5		80052477 (1889, 1890) Novel Protein sim. GBank gil732353 splP39606 YWCH_BACSU - HYPOTHETICAL 36.6 KD PROTEIN IN QOXD-VPR INTERGENIC REGION	Contains protein domain (PF00818) - UNCLASSIFIED Ice nucleation protein repeat	UNCLASSIFIED	264906, 264604, 264605, 265020, 18108387	
946	79248402 (1891, 1892)				265017	

947	81802699 (1893, 1894)	81802699 (1893, 1894) Novel Protein sim. GBank gi[2896770 emb CAA17247 - ((AL021899) hypothetical protein Rv2033c [Mycobacterium tuberculosis]	Contains protein domain (PF00459) - phosphatase Inositol monophosphatase family	phosphatase	18108394, 22278996, 264907, 264909, 265008, 265009, 264910, 264758, 264600, 264602, 265018, 264605, 264769, 264689, 264603,
948	88165538 (1895, 1896)	88165538 (1895, 1899) Novel Protein sim. GBank gil2827284 (AF041037) - novel antagonist of FGF signaling [Homo sapiens]		igi	18108398, 56182575, 22278997, 22278999, 60432049, 29331822, 29331826, 264907, 56182435, 55811386, 265011, 264600, 265017, 265018, 18108351, 265020, 265021, 265022, 27486265, 263972, 55811576, 264638, 60170394, 264566
949	88081786 (1897, 1899) Novel Protein sim. gil4507985[ref]NP_ (clone pHZ-17)	GBank 003427.1 pZNF1 - zinc finger protein 135	Contains protein domain (PF00096) - Iranscriptfactor Zinc finger, C2H2 type		29331825, 21906764, 27486261, 21906766, 52644296, 33657349, 87168518, 56994075, 265020, 265021, 87168559, 52644150, 264637
920	79485872 (1899, 1900)		Contains protein domain (PF00038) - struct Intermediate filament proteins		264683, 18108361
951	20451411 (1901, 1902)	20451411 (1901, 1902) Novel Protein sim. GBank gil5420387[emb CAB45679.1 - (AJ243459) proteophosphoglycan (Leishmania major)		SIFIED	264604
952	79566954 (1903, 1904)) Novel Protein sim. GBank gi 5305702[gb AAD41779.1 AF12686 - (AF126867) calpain- like protease [Mus musculus]		cathepsin	264910, 264691
953	10196003 (1905, 1906)	$\overline{}$		transport	264510
954	9893326 (1907, 1908)	Novel Protein sim. GBank gi 2360965 (AF016253) - D- amino acid dehydrogenase [Klebsiella aerogenes]		dehydrogenase	264508
955	95313410 (1909, 1910)	Bank 06319.1pSIPI - SYT interacting protein	Contains protein domain (PF00076) - dna_rna_bind RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_rna_bind	56994075, 264509, 264905, 264906, 264907, 264908, 264909, 264910, 264758, 264760, 264758, 264762, 264761, 264760, 18108351, 264767, 264764, 264766, 26468, 264767, 264687, 264768, 264769, 264689, 264629, 264639, 264631, 264631, 264633, 264639, 18108388, 264563, 264564
926	80064224 (1911, 1912) Novel Protein sim. G (294752) rimJ [Myco) Novel Protein sim. GBank gi 2052129 emb CAB08155 - (294752) rimJ [Mycobacterium tuberculosis]			264605
957	80056206 (1913, 1914	80056206 (1913, 1914)		UNCLASSIFIED	264603, 18108362
958	80036446 (1915, 1916)) Novel Protein sim. GBank gi 1709767 sp Q00451 PRF1_LYCES - 36.4 KD PROLINE- RICH PROTEIN			264908, 264910, 264762, 263978, 264637
656	80026647 (1917, 1918	80026647 (1917, 1918) Novel Protein sim. GBank gil2131050jemb CAB09260 - (Z95844) opcA [Mycobacterium tuberculosis]		UNCLASSIFIED	264602, 264692
096	37815406 (1919, 1920)	37815406 (1919, 1920) Novel Protein sim. GBank gil2129478 pirl S51939 - Chitinase (EC 3.2.1.14) precursor - beet		UNCLASSIFIED	264259
961	20567383 (1921, 1922)				263978
985	11399318 (1923, 1924)				264593

	80590374 (1925, 1926)			UNCLASSIFIED	264510, 264288, 264555, 264556, 264559, 264486
964	79832019 (1927, 1928) Novel Protein sim. GE (AB023206) KIAA098	Novel Protein sim. GBank gil4589622[db] BAA76833.1] - (AB023206) KIAA0989 protein [Homo sapiens]		UNCLASSIFIED	264112, 264910, 264689
965	91229485 (1929, 1930)	91229485 (1929, 1930) Novel Protein sim. GBank gi[5420387 emb[CAB46579.1] - (AJ243459) proteophosphogiycan [Leishmania major]		UNCLASSIFIED	264488, 265017, 264448, 264634, 264558, 83373044
996	95292815 (1931, 1932)			UNCLASSIFIED	264906, 264592, 264596, 264604, 264768,
					Z1905/04, Z0489Z, Z64893, Z046Z9, Z04636, 264638
96	79255708 (1933, 1934) Novel Protein sim. Gl	Novel Protein sim. GBank		phosphatase	264760
		TRANSDUCTION PROTEIN REGX3	Response regulator receiver domain		
998	79560269 (1935, 1936)	79560269 (1935, 1936) Novel Protein sim. GBank gij2661836jemb]CAA75187j - I/Y14964) nutativa transport protein Mathulpohitus		transport	264693
	methylotrophus	methylotrophus]			
696	79919470 (1937, 1938)	Novel Protein sim. GBank gi[5419878 emb CAB46422.1] - (AL096747) hypothetical protein [Homo sapiens]	Contains protein domain (PF00096) - dna_ma_bind Zinc finger, C2H2 type	dna_ma_bind	35696286, 264685, 264686, 35695917, 264692, 18108374, 264635
970	95085947 (1939, 1940)			UNCLASSIFIED	18108392, 18108394, 18108398, 22278995,
					22278996, 22278998, 22278999, 29147620,
					264628, 2650006, 265001, 2650008, 2650009,
					18108348, 33109954, 2550110, 255011,
					10100331, 204200, 21900/67, 21900/06,
					18108374, 18108374, 18108377, 28463U,
					204033, 18106380, 83373044, 18106367, 18108388
971	78919770 (1941, 1942)			UNCLASSIFIED	265007, 265020, 22279002
972	20710704 (1943, 1944)				264557
973	20370183 (1945, 1946) Novel Protein sim. Gl	Novel Protein sim. GBank			264604
		gi 1723119 sp P53990 Y174_HUMAN - HYPOTHETICAL PROTEIN KIAA0174			
974				UNCLASSIFIED	264565
975				UNCLASSIFIED	264510
976	80205742 (1951, 1952) Novel Protein sim. G	Novel Protein sim. GBank gij3881459jembjCAA92988.1 -		UNCLASSIFIED	264508, 264906, 264758, 264632, 264639,
		(268753) predicted using Genefinder; Similarity to Yeast			264563
		hypothetical protein YIK9 (SW/YIK9_YEAST); cDNA EST			
		EMBL-D27879 comes from this gene, conv. Co.			
		EMBL: D64477 comes from this ge			
977	10355349 (1953, 1954)	10355349 (1953, 1954) Novel Protein sim. GBank		UNCLASSIFIED	264906
		gil549456 sp Q05335 XYS3_PSEPU - XYLDLEGF			
		OPERON TRANSCRIPTIONAL ACTIVATOR 3			
88	80025927 (1955, 1956)			UNCLASSIFIED	264600, 264602, 264603, 264604
979	80447820 (1957, 1958)	Novel Protein sim. GBank gi[3171904 emb CAA75869 - (Y15908) DIA-12C protein [Homo sapiens]		UNCLASSIFIED	264767, 264768, 265006, 265007, 264906
980	80025928 (1959, 1960)	+		UNCLASSIFIED	264600, 264602, 264605
981	80098550 (1961, 1962) Novel Protein sim. Gl	Rovel Protein sim. GBank gi[359940 (AF017368) -		UNCLASSIFIED	264692, 264555, 264556, 264557, 264559
		inducigeninal dyspinasia protein z jimus musculus			

080	80105670 (1063 1064)	80195670 (1963 1964) Novel Protein eim GRank nit2950220 lemblCAA715751.		UNCLASSIFIED	264404
1		(Y10545) fused-ccdB (Escherichia coli)			
983	90995041 (1965, 1966)	90995041 (1965, 1966) Novel Protein sim. GBank gi 476389 pir B43402 - myosin		struct	65274572, 56182575, 264908, 264909,
		heavy chain-B, neuronal - chicken			265007, 265008, 264758, 265010, 55811150, 33657023, 264634, 264557, 264558
984	20466876 (1967, 1968)			UNCLASSIFIED	264605
985	65461368 (1969, 1970) Novel Protein sim. (AJ007747) hypoth bronchiseptical	Novel Protein sim. GBank gij2451504 emb CAA07650.1 - (AJ007747) hypothetical protein BbLPS1.21 [Bordetella bronchiseptical	Contains protein domain (PF00534) - transferase Glycosyl transferases group 1	transferase	56182435, 264600
986	87102868 (1971, 1972)	-		UNCLASSIFIED	264106, 264110, 265020, 60170615
987	79867231 (1973, 1974)				264909
988	19858661 (1975, 1976)			UNCLASSIFIED	264600
686	88095329 (1977, 1978)			UNCLASSIFIED	264508, 265017, 264534, 264564
066	88057746 (1979, 1980) Novel Protein sim.	Novel Protein sim. GBank	Contains protein domain (PF01421) - oxidase	oxidase	264259, 264908, 265009, 264910, 264596,
		gi 5725506 gb AAD48080.1 AF06015 - (AF060152) METH1 Reprolysin (M12B) family zinc protein IHomo sapiens	Reprolysin (M12B) family zinc metalloprotease		264369, 264288, 264766, 264628, 264635, 264566
66	10106140 (1981 1982)			UNCLASSIFIED	264909
5 5	70846604 (4082, 4084) Novel Drotein eim	Novel Destein sim CBank ail9405040lembICAB088351	Contains profein domain (PE00211) - HINCH ASSIEIED	INCI ASSIFIED	26450R 264593
788	18040094 (1803, 1804)		Adenylate and Guanylate cyclase		2001
		Iuberculosis	catalytic domain		
993	10814053 (1985, 1986)	-			264907
į	440000000000000000000000000000000000000	11 C		operation pro-	SEARINS
66 76	11090590 (1987, 1988) 	11090590 (1987, 1988) Novel Protein sim. Gbank gij3329297 (AE001355) - Ribonucleoside Reductase, Large Chain [Chlamydia Itrachomatis]		reductase	2040UZ
908	94321911 (1989 1990	94321911 (1989, 1990) Novel Protein sim GBank	Contains protein domain (PF00176) -	helicase	18108398, 65274572, 22278996, 264490,
\$66 C	94321911 (1989, 1990)	Jilovei Protein sim. GBank gil5106572[gb]AD39760.1 AF14394 - (AF143946) transcriptional activator SRCAP [Homo sapiens]	SNF2 and others N-terminal domain Contains protein domain (PEDDECK) through alace	nelicase	16106396, 032.49.77, 222.70939, 204430, 20403049, 2933.18.27, 29146498, 264508, 264032049, 264508, 264591, 264591, 264591, 264591, 264592, 2643768, 21906754, 265010, 265011, 265018, 264068, 21906765, 21906768, 21906769, 55811957, 264692, 264693, 264629, 35696423, 55811576, 3569685, 264556, 264556, 264556, 264556, 264556, 264556, 264556, 264556, 264556, 264556, 264568, 293373044, 22279002, 264558, 264556, 264558, 264558, 264568, 293373044, 22279002, 264568, 264556, 264556, 264558, 264588, 2
9 8 8	91013/45 (1991, 1992) Novei Protein sim. KIA001LB (Homo), Nover Frotein sim. Gaank gijzs 11719 (ACUU4227) - KIA001LB [Homo sapiens]	Contains protein domain (Froossy) - PDZ domain (Also known as DHR or GLGF).	pilospijatase	55526486, 87168518, 264910, 264906, 264565, 264566, 264693, 264766
997	80503347 (1993, 1994) Novel Protein sim. Iransporter, ATP-bi	Novel Protein sim. GBank gi[2649101 (AE001001) - ABC transporter, ATP-binding protein [Archaeoglobus fulgidus]	Contains protein domain (PF00005) - transport ABC transporter	transport	35696286, 22278997, 22278999, 264508, 264905, 264908, 265010, 264600, 264602,
					264605, 264688, 264769, 265021, 264565, 18108391
866	11397390 (1995, 1996) Novel Protein sim.				264595
		gij123530 sp P04929 HRPX_PLALO - HISTIDINE-RICH GLYCOPROTEIN PRECURSOR			
	4				

264682	264602	60432049, 264907, 264909, 264511, 264603, 264683, 264683, 264689, 29148627, 21906769, 264692, 18108385, 22279000	265009, 264369, 265020	265019	264635	264508	264259	264905	265007, 264602, 264605, 264760, 264636	264102, 264288	264592	35696052, 264905, 264764, 264768, 35695917, 264629	35696052, 264905, 264600, 264601, 264602, 264605, 264762, 264766, 264768, 264689	264591, 21906768	29331824, 265019, 265020
UNCLASSIFIED	UNCLASSIFIED	kinase	struct			UNCLASSIFIED	UNCLASSIFIED	synthase	synthase	kinase	dehydrogenase	UNCLASSIFIED	phosphorylase	UNCLASSIFIED	
		Contains protein domain (PF00023) - kinase Ank repeat								Contains protein domain (PF00595) - kinase PDZ domain (Also known as DHR or GLGF).			Contains protein domain (PF00013) - phosphorylase KH domain		
Novel Protein sim. GBank gil25068971spIP46490 YFCA_HAEIN - HYPOTHETICAL PROTEIN HI0198		GBank gil2224699 dbj BAA20833 - 379 [Homo sapiens]	Novel Protein sim. GBank gij586121 sp[P37709 TRHY_RABIT - TRICHOHYALIN		Novel Protein sim. GBank giļ854065 emb CAĀ58337 - (X83413) U88 [Human herpesvirus 6]	Novel Protein sim. GBank gi[114073]sp P07672 APT_ECOL - ADENINE		Novel Protein sim. GBank gi 4062979 dbj BAA36210.1 - (AB017138) epsilon subunit of malonate decarboxylase [Pseudomonas putida]	Novei Protein sim. GBank gi 2808807 emb CAA04607.1 - (AJ001206) putative trehalose synthase Streptomyces coelicolor]	GBank gij3639077 (AF090113) - AMPA otein [Rattus norvegicus]	GBank gil3746332 (AF016307) - possible oxidase, may function as a demethylase lilotij	Novel Protein sim. GBank gi[2342647]gb AAB86591.1 - (U90653) DHHC-domain-containing cysteine-rich protein [Homo sapiens]	GBank gij3413411[emb CAA20272] - sine pentaphosphate synthetase/ nucleotidyltransferase [Streptomyces		
11768047 (1997, 1998) Novel Protein sim. gil2506697 sp P46- PROTEIN HI0198	20727944 (1999, 2000)	86673131 (2001, 2002) Novel Protein sim. (AB002377) KIAA0	80189603 (2003, 2004) Novel Protein sim. gil586121 sp P377	17933491 (2005, 2006)	16314987 (2007, 2008) Novel Protein sim. (X83413) U88 [Hur	79617144 (2009, 2010)	37815429 (2011, 2012)	79620871 (2013, 2014) Novel Protein sim. (AB017138) epsilo [Pseudomonas pul	88094444 (2015, 2016)	57451289 (2017, 2018)	94672537 (2019, 2020)	85546916 (2021, 2022) Novel Protein sim. (U90653) DHHC-d [Homo sapiens]	95294456 (2023, 2024) Novel Protein sim. (AL031231) guano. polyribonucleotide coelicolor)	86095772 (2025, 2026)	1014 86608828 (2027, 2028)
666		1001	1002	1003	1004	1005		1007	1008	1009	1010	1011	1012	1013	1014

22278994, 22278995, 56994075, 22278996, 22278999, 264259, 29331825, 29331828, 264907, 56182435, 264510, 264591, 264997, 56182435, 264510, 264591, 264997, 265044, 55812038, 294758, 21906754, 33657084, 265010, 264600, 265017, 265018, 265019, 18108351, 21906765, 21906766, 21906767, 21906768, 233657023, 66274620, 33657182, 23857023, 65274620, 33657182, 32833986, 18108370, 18108377, 55811576, 35696423, 264630, 22279000, 264565	264686, 264693	264600	22278996, 29148627, 264563	264686	264508, 264509, 264512, 264600, 264762, 264769, 264689, 18108370, 264636, 264638, 264486	264769	35696052, 264508, 265008, 265009, 264769, 18108387, 264563	264593	52646842, 33657402, 33657023, 18108379, 55811576, 264631, 264556, 264557, 264559, 18108388, 264566	264693	264602	265017
struct	esterase	transport		UNCLASSIFIED		synthase	phosphatase	- ATPase_associated	UNCLASSIFIED	transport	helicase	UNCLASSIFIED
Contains protein domain (PF00040) - struct Fibronectin type II domain		Contains protein domain (PF00496) - transport Bacterial extracellular solute-binding proteins, family 5					Contains protein domain (PF00459) - phosphatase Inositol monophosphatase family	Contains protein domain (PF00122) - ATPase_associated 264593 E1-E2 ATPase				
1015 95418879 (2029, 2030) Novel Protein sim. GBank gil4159995 (AF083095) - SELIL. [Mus musculus]	2) Novel Protein sim. GBank gi[2506969]sp P41407 ACPD_ECOLI - ACYL CARRIER PROTEIN PHOSPHODIESTERASE (ACP PHOSPHODIESTERASE)	11069213 (2033, 2034) Novel Protein sim. GBank gi 5103943 dbj BAA79259.1 - (AP000059) 602aa long hypothetical oligopeptide-binding protein oppA [Aeropyrum pernix]			0	37036243 (2041, 2042) Novel Protein sim. GBank gil4633807lgb AAD26859.1 AF12779 - (AF127795) trehalose biosynthetic enzyme TreY (Rhizobium leguminosarum bv. viciae]	80502627 (2043, 2044) Novel Protein sim. GBank gij1781230 emb CAB06277 - (283867) hypothetical protein Rv3137 [Mycobacterium tuberculosis]	11399341 (2045, 2046) Novel Protein sim. GBank gi 3777495 (U92083) - calcium Itransporting ATPase [Pichia angusta]		79644200 (2049, 2050) Novel Protein sim. GBank gij3483045 emb CAA20556 - (AL031371) putative transport system permease protein	80025946 (2051, 2052) Novel Protein sim. GBank gil1174922 sp Q02322UVRD HAEIN - DNA HELICASE II	17659234 (2053, 2054) Novel Protein sim. GBank gil4757728 ref NP_004886.1 pAGTA - angiotensin/vasopressin receptor AlIAVP-like
95418879 (2029, 2030	79559694 (2031, 2032)	11069213 (2033, 2034	80072430 (2035, 2036	11703607 (2037, 2038	80234432 (2039, 2040)	37036243 (2041, 2042	80502627 (2043, 2044	11399341 (2045, 2046	80057129 (2047, 2048)	79644200 (2049, 2050	80025946 (2051, 2052	17659234 (2053, 2054
1015		1017	1018		1020	1021	1022	1023	1024	1025	1026	1027

028	1028 20297928 (2055, 2056) Novel Protein sim.	Novel Protein sim. GBank gi[2791409 emb CAA16003 -	Contains protein domain (PF00330) - UNCLASSIFIED	UNCLASSIFIED	264600
		(AL021184) acn [N	Aconitase family (aconitate hydratase)		
1029	94665090 (2057, 2058)			UNCLASSIFIED	264595
1030	88095343 (2059, 2060) 			UNCLASSIFIED	264907, 264908, 264510, 264512, 265009, 265010, 265011, 264600, 264602, 264603, 264605, 26473, 18108374
1031	95289117 (2061, 2062)			UNCLASSIFIED	264905, 264906, 264909, 264595, 264692, 264630, 264634, 264638
1032		Novel Protein sim. GBank gil4503895[ref]NP_000145.1 pGALK - galactokinase 1		UNCLASSIFIED	264689
1033	86464818 (2065, 2066) Novel Protein sim.			UNCLASSIFIED	35696052, 264906, 264510, 18108354,
		hypothetical protein [Aquifex aeolicus]			264687, 264769, 264689, 60431602, 18108385, 264486
	79245937 (2067, 2068)	Novel Protein sim. GBank gil405895 (U00007) - methionyl- tRNA synthetase [Escherichia coli]		UNCLASSIFIED	264906
1035	79956355 (2069, 2070)			UNCLASSIFIED	264692
	85804998 (2071, 2072)				264905, 66712502, 264908, 264766
_	87896058 (2073, 2074)			UNCLASSIFIED	29331824, 264909, 60433438, 265019
1038	20481015 (2075, 2076) Novel Protein sim. kidney disease-ass	Novel Protein sim. GBank gil790819 (L39891) - polycystic kidney disease-associated protein [Homo sapiens]	Contains protein domain (PF01477) - PLAT/LH2 domain		264604, 264634
1039	87260021 (2077, 2078) Novel Protein sim. IFauine heroesviru	Novel Protein sim. GBank gi[2605967 (AF030027) - 24 Fauine heroesvirus 4]		UNCLASSIFIED	264092, 264093, 264094, 264683, 264689,
1040	80026840 (2079, 2080)	80026840 (2079, 2080) Novel Protein sim. GBank gil2352095 (U97022) - DNA topolsomerase I Fervidobacterium Islandicum)	Contains protein domain (PF01131) - isomerase Prokarvotic DNA tonoisomerase	isomerase	264595
1041		10156682 (2081, 2082) Novel Protein sim. GBank gil3256535 dbj BAA29218.1 - (AP000001) 301aa long hypothetical 2-phosphoglycerate		kinase	264907
\$55	11084375 /2083 2084)	Marial Destriction Charles 1100 600001			
	1004373 (2003, 2004)	(X99309) Action (X99309) ARI protein [Drosophila melanogaster]			264605
	80057136 (2085, 2086)	Novel Protein sim. GBank gi 1870167 emb CAA70125 - (Y08921) msiK [Streptomyces reticuli]	Contains protein domain (PF00005) - Irransport ABC transporter	transport	264565, 264567
1044	80025952 (2087, 2088)			UNCLASSIFIED	265006, 264602, 265017
	52415482 (2089, 2090) Novel Protein sim. (AL109732) hypotl A3(2)]	Novel Protein sim. GBank gijs689890 emb CAB52053.1] - (AL109732) hypothetical protein [Streptomyces coelicolor A3(2)]			29331825, 264637
	11754862 (2091, 2092)	Novel Protein sim. GBank gil854065 emb CAA58337 - (X83413) U88 (Human herpesvirus 6)		helicase	264686
	37036258 (2093, 2094) Novel Protein sim. (AB019033) orfSA	Novel Protein sim. GBank gil4210471 dbj BAA74535.1 - (AB019033) orfSA [Pseudomonas sp.]		UNCLASSIFIED	264769
	79186400 (2095, 2096) Novel Protein sim. (AL031232) hypotl coelicolor]		Contains protein domain (PF00220) - UNCLASSIFIED Neurohypophysial hormones, N-terminal Domain	UNCLASSIFIED	264687
	81755108 (2097, 2098) Novel Protein sim. gi 5051636 gb AA(domain-binding mi	Novel Protein sim. GBank gi 5051636 gb AAD38326.1 AF07372 - (AF073727) EH domain-binding milotic phosphoprotein [Homo sapiens]		UNCLASSIFIED	264905, 264634
1050	79471521 (2099, 2100)			UNCLASSIFIED	264686

1051	180475471 (2101 2102)			UNIO A COLDIED	10100374 264760 265010 265011 264601
	7-0-1-0-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-			ONCE ASSESSED IN THE PARTY OF T	265009, 264604, 264605, 264636, 18108351,
	_				264692
1052	82442962 (2103, 2104)	82442962 (2103, 2104) Novel Protein sim. GBank	Contains protein domain (PF00389) - dehydrogenase	dehydrogenase	264508, 264762, 264687, 264486
		gi 3123275 sp P35136 SERA_BACSU - D-3-	D-isomer specific 2-hydroxyacid		
		PHOSPHOGLYCERATE DEHYDROGENASE (PGDH)	dehydrogenases		
1053		94851640 (2105, 2106) Novel Protein sim. GBank gij5441319 emb CAB46717.1 -			264686, 18108374, 29331824, 83373044,
		(AL034396) dJ1158B12.1 (zinc finger, X-linked, duplicated			21906754, 52645156, 56182435, 264689,
		A) [Homo sapiens]			29331827, 27486261, 35696052, 21906765,
					35696423, 21906768, 56182575, 21906769,
					55811957, 87168518, 35696286, 22278997,
					265020, 265011, 265021, 265022, 265007,
					265018, 22279000, 22279002, 264482,
					264906, 52644150, 264909, 264288,
					29331822, 52645080, 264766
1054				UNCLASSIFIED	264686
1055	80594138 (2109, 2110)	80594138 (2109, 2110) Novel Protein sim. GBank	Contains protein domain (PF00270) - helicase	helicase	264907, 264602, 264681, 264288, 21906768,
		gi[5052508 gb AAD38584.1 AF14560 - (AF145609)	DEAD/DEAH box helicase		33657109, 55810764, 35695855, 264631
	-	BcDNA.GH02833 [Drosophila melanogaster]			
1056		17882319 (2111, 2112) Novel Protein sim. GBank gij3021676 dbj BAA25358 -		rnapolymerase	264906
		(D86033) RNA polymerase sigma-70 factor (Pseudomonas			
		fluorescens)			
1057		85667216 (2113, 2114) Novel Protein sim. GBank ail 1226281 (U50308) - No		UNCLASSIFIED	264682
		definition line found [Caenorhabditis elegans]			
1058					264764
1059		94662754 (2117, 2118) Novel Protein sim. GBank		transcriptfactor	35696052, 35695855, 265009, 264636
		ail1170016isolP46808iGREA MYCLE - TRANSCRIPTION		-	
_		ELONGATION FACTOR GREA (TRANSCRIPT CLEAVAGE			
		FACTOR GREA)			
1060	79481169 (2119, 2120)	Novel Protein sim. GBank		glycoprotein	29146499, 264681, 264683, 264687
		gi[2499087]sp Q09332 UGGG_DROME - UDP-			
		GLUCOSE: GLYCOPROTEIN GLUCOSYLTRANSFERASE			
		PRECURSOR (DUGT)			
1061		11034025 (2121, 2122) Novel Protein sim. GBank gij90254 pir JA28334 - protein-		phosphatase	264634
		lyrosine-phosphatase (EC 3.1.3.48) Ly-5 precursor (B-cell			
		variant) - mouse			
1062		39567937 (2123, 2124) Novel Protein sim. GBank		dehydrogenase	264593
		gij3334200jspj049954jGCSP_SOLTU - GLYCINE			
		DEHYDROGENASE (DECARBOXYLATING) PRECURSOR			
		(GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE			
		SYSTEM P-PROTEIN)			
1063	8490481 (2125, 2126)	Novel Protein sim. GBank			264508
		gi[2499966 sp Q41228 PSE1_NICSY - PHOTOSYSTEM I			
		NEACTION CENTINE SUBDINITION A PRECURSOR (PSI-E			
1064		78891783 (2127, 2128) Novel Protein sim GBank dil82654 Initi LADD86 - 10K zein			265007 265008 18108351 18108385
<u>:</u>		precursor - maize			200001 1000001 200001

1065	80021208 (2129, 2130)	Novel Protein sim. GBank gil2120998 pir S70682 -		transferase	264600, 264602, 264689
900	47006070 (2424 2422)	glycosymansierase nomonog - por detenia per ussis			000330
90	17690679 (2131, 2132)	1/0900/3 (2131, 2132) Novel Protein Sim. GBank gip506362lsp P15042 DNLJ_ECOLI - DNA LIGASE		synmase	\$00coz
		(POLYDEOXYRIBONUCLEOTIDE SYNTHASE (NAD+))			
1067	_				264909
1068		Bank gil4007669 emb CAA22355 - oxidoreductase [Streptomyces	Contains protein domain (PF00248) - reductase Aldo/keto reductase family		264688, 18108362, 264558, 264600, 264760
1069	83002954 (2137, 2138) Novel Protein sim. Gi (AB023143) KIAA092	Novel Protein sim. GBank gil4589484[dbj[BA476770.1] - (AB023143) KIAA0928 protein [Homo sapiens]		UNCLASSIFIED	264604
1070	82101992 (2139, 2140)	82101992 (2139, 2140) Novel Protein sim. GBank		UNCLASSIFIED	264604, 264760
		gij 20304(spjP15932jFLGK_SAL1Y - FLAGELLAR HOOK- ASSOCIATED PROTEIN 1 (HAP1)			
1071	20710589 (2141, 2142)	Novel Protein sim. GBank gi 1750397 (U81261) - glutamate synthase large subunit [Pseudomonas aeruginosa]		synthase	264602
1072	1072 82356540 (2143, 2144)			UNCLASSIFIED	264687, 264688, 21906764, 35696052,
					35895917, 35895855, 264600, 264601, 264602, 265009, 264605, 264508, 264905,
573	70814400 (2145 2145)				264690, 264906, 264762, 264628, 264766
	13014400 (2143, 2140)				20203
1074	1074 80105992 (2147, 2148) Novel Protein sim. Gl protein - mouse (frag	Bank gi 477532 pir A49175 - Motch B ment)	Contains protein domain (PF00008) - synthase EGF-like domain	synthase	264906
1075		81850293 (2149, 2150) Novel Protein sim. GBank gij3893109jemb CAA76940 -		UNCLASSIFIED	56994075, 22278998, 264594, 264757,
		(Y17920) CALO protein [Drosophila melanogaster]		,	264596, 265018, 265019, 264681, 264369, 264688, 265020, 18108364, 18108374
1076	80477264 (2151, 2152)	80477264 (2151, 2152) Novel Protein sim. GBank		ATPase_associated 264769	264769
	-	gi 1176203 sp P46442 YHCM_ECOLI - HYPOTHETICAL 43.1 KD PROTEIN IN RPLM•HHOA INTERGENIC REGION (F375)			
1077	79831334 (2153, 2154)	79831334 (2153, 2154) Novel Protein sim. GBank gi 4033487 sp Q44472 TUD4_AGRVI - PUTATIVE HYDROXYPYRUVATE REDUCTASE			264905
1078					264600
1079		80494518 (2157, 2158) Novel Protein sim. GBank gi 3413828 emb CAA20296 - (AL031260) hypothetical protein SC9A10.09 [Streptomyces coelicolor]		UNCLASSIFIED	18108394, 264769, 264634, 264636
1080	11767188 (2159, 2160)			UNCLASSIFIED	264684
1081	94747080 (2161, 2162)				83373044, 265019, 22279002, 264482, 18108351, 264682, 264908, 264693, 264487
1082	81490656 (2163, 2164)			UNCLASSIFIED	264758, 264768, 264769, 21906767, 264511,
					264910, 264634, 264635, 264905, 264636, 264906, 264637, 264907, 264908, 264764,
					264638, 20281099, 264766, 264595

60424179, 264905, 264906, 264510, 6043229, 264759, 87168474, 264605, 264769, 264689, 18108364, 18108376, 35695855, 264636	264769	60432229, 264687	264486	29331827, 264693	264905, 264601, 18108387	264488, 264511, 265011, 264682, 264768, 264689, 21906764, 35695917, 265020, 32833986, 18108370, 35695855	264604	264637	264566	264769, 35695917, 35695855, 254600, 264602, 264603, 264605, 18108351	264686	264906, 265007, 264595, 264600, 264602, 264603, 264604, 264605, 264762, 264766, 264769, 264636, 26458, 18108387, 60432113, 264482, 264486	264600, 264602, 264603, 264604, 264605, 35695917, 264692, 264631	265019	264687
UNCLASSIFIED	UNCLASSIFIED	eph	ribosomalprot	helicase	UNCLASSIFIED	UNCLASSIFIED		UNCLASSIFIED		nuclease	UNCLASSIFIED	nuclease	synthase	UNCLASSIFIED	synthase
		Contains protein domain (PF00118) - eph TCP-1/cpn60 chaperonin family	Contains protein domain (PF00238) - inbosomalprot Ribosomal protein L14	Contains protein domain (PF00385) - chromo' (CHRromatin Organization MOdifier) domain		Contains protein domain (PF00329) - UNCLASSIFIED Respiratory-chain NADH dehydrogenase, 30 Kd subunit				Contains protein domain (PF00730) - nuclease Endonuclease III		Contains protein domain (PF00455) - nuclease Bacterial regulatory proteins, deoR family			
	Novel Protein sim. GBank gij418384[sp]P32057[WCAI_ECOLI - PUTATIVE COLANIC ACID BIOSYNTHESIS GLYCOSYL TRANSFERASE WCAI	Novel Protein sim. GBank gil 1899190 (U90204) - heat shock protein 60 [Tsukamurella tyrosinosolvens]		Novel Protein sim. GBank gil4160198 emb CA415431 - Contains protein domain (PF00385) (AL008583) dJ327J16.3 (novel CHROMObox family protein) 'chromo' (CHRromatin Organization [Homo sapiens]	Novel Protein sim. GBank gil2983155 (AE000693) - phosphoglucomutase/phosphomannomutase (Aquifex aeolicus)		11083825 (2179, 2180) Novel Protein sim. GBank gi 4007680 emb CAA22366 - (AL034443) putative oxidoreductase [Streptomyces coelicolor]	Novel Protein sim. GBank gi[2495562 sp P77239 YLCD_ECOLI - HYPOTHETICAL 44.3 KD PROTEIN IN NFRB-PHEP INTERGENIC REGION PRECURSOR		GBank gi[2960098]emb CAA17996.1] - ycobacterium tuberculosis]	GBank gil1001642 dbj BAA10373 - phosphohydrolase [Synechocystis sp.]	Novel Protein sim. GBank gild 585587 emb CAB40855.1 . (AL049628) putative adenine glycosylase [Streptomyces coelicolor]	Novel Protein sim. GBank gi 115001 sp P19206 BIOB_BACSH - BIOTIN SYNTHASE (BIOTIN SYNTHETASE)		Novel Protein sim. GBank gi 114135 sp P08205 ARGA_ECOLI - AMINO-ACID ACETYLTRANSFERASE (N-ACETYLGLUTAMATE SYNTHASE) (AGS)
1083 87446717 (2165, 2166) Novel Protein sim. gil1722945 sp Q10 44.6 KD PROTEIN	37799306 (2167, 2168) Novel Protein sim. gil418384 sp P320 ACID BIOSYNTHE	86475368 (2169, 2170) Novel Protein sim. Shock protein 60 [7]	79608269 (2171, 2172)	79603979 (2173, 2174) Novel Protein sim. (AL008583) dJ327 (Homo sapiens)	79854963 (2175, 2176) Novel Protein sim. phosphoglucomuta aeolicus]	80216800 (2177, 2178) Novel Protein sim. gil4981768 gb AAC dehydrogenase, 3C maritima]	11083825 (2179, 2180) Novel Protein sim. (AL034443) putativ coelicolor]	12917471 (2181, 2182)	80252286 (2183, 2184)	80496304 (2185, 2186) Novel Protein sim. (AL022121) nth [M	10880972 (2187, 2188)	87457250 (2189, 2190) Novel Protein sim. (AL049628) putativ coelicolor]	80025977 (2191, 2192) Novel Protein sim. gij115001 sp P192 (BIOTIN SYNTHE'	79239560 (2193, 2194)	
1083	1084	_	1086	1087					1092					1097	

100 657-6671 230	9	20100 20100 000000				
Move Protein sim. GBank Move Protein sim. GBank g M	200	38353030 (Z.191, Z.190)) Novel Protein sim. Gbank gijsg16144spg033017tmD_MYCLE - TRNA (GUANINE- NI)-METHYLTRANSFERASE (M1G- METHYLTRANSFERASE) (TRNA (GM37) METHYLTRANSFERASE)			264603
DESATURASE (STEATON-COA	<u>1</u>	(85736571 (2199, 2200)	Novel Protein sim. GBank		desaturase	264259, 264636
ACID DESATURASE) (DELTA(9)-DESATURASE)			gij3023255Isp Q64420 ACOD_MESAU - ACYL-COA DESATIJRASF (STEAROY) -COA DESATIJRASE) (FEATTY			
Contains protein domain (PD1336) Isomerase			ACID DESATURASE) (DELTA(9)-DESATURASE)			
TOPOISOMERASE LONGGA-PACTIEN (RELAXING Ingermerase DNA binding C4 zinc 10POISOMERASE LONGGA-PACTIEN (RELAXING Ingermerase DNA binding C4 zinc 10POISOMERASE LONGGA-PACTIEN (RELAXING Ingermerase DNA binding C4 zinc ENZYME (UNIVMSTING ENZYME (SWIVELASE) 10	101	80491857 (2201, 2202)	Novel Protein sim. GBank	Contains protein domain (PF01396) -	isomerase	264769
TOPOISOMERANSE CONCIONERANCE CONCIONERANCE			gi 1174735 sp P43012 TOP1_HAEIN - DNA	Topoisomerase DNA binding C4 zinc		
198777514 (2202, 2204) World Protein sim, GBank gil 1908596 (U81788) - kinesin-73 Struct			TOPOISOMERASE I (OMEGA-PROTEIN) (RELAXING ENZYME) (UNTWISTING ENZYME) (SWIVELASE)	linger		
B1897259 (2205, 2206) Navel Protein sim. GBank gil1906596 (U61788) - kinesin-73 struct	1102				UNCLASSIFIED	264910, 264909
Drosophila melanogaster 95003115 (2207, 2208) Novel Protein sim. GBank gil2935448 (AF048976) - Synaptic ras GTPase-activating protein p135 SynGAP 80255121 (2209, 2210) Synaptic ras GTPase-activating protein p135 SynGAP Rattus novegicus Rattus nov	1103				struct	264757
95003115 (2207, 2208) Novel Protein sim. GBank gilz835448 (AF0489*6) - Synaptic ras GTPase-activating protein p135 SynGAP Rattus norvegicus Rattus norvegicus Rattus norvegicus		\neg	[Drosophila melanogaster]			
Rattus norvegicus Rattus norvegicus Rattus norvegicus Rattus norvegicus Rattus norvegicus Rattus norvegicus Rattus norvegicus Rattus norvegicus Rattus	102		Novel Protein sim. GBank gi 2935448 (AF048976) -		UNCLASSIFIED	29331822, 21906754, 264555, 264556,
Raffus norvegicus Raffus norvegicus Raffus norvegicus 190256121 (2209, 2210) 19034101 (2211, 2212) 190440616 (2215, 2216) 190440616 (2215, 2216) 190440616 (2215, 2216) 190440616 (2215, 2216) 190440616 (2215, 2216) 190440616 (2215, 2216) 190440616 (2215, 2216) 190440616 (2215, 2216) 190440616 (2217, 2218) 19044061			synaptic ras GTPase-activating protein p135 SynGAP			264558, 22279002
80255121 (2209, 2210) 8040516 (2213, 2214) 80440616 (2215, 2216) Novel Protein sim. GBank 80440616 (2215, 2216) Novel Protein sim. GBank 801173421sp P43416 SECY_STRSC - PREPROTEIN TRANSLOCASE SECY SUBUNIT TRANSLOCASE SECY SUBUNIT TRANSLOCASE SECY SUBUNIT Relicase Steptomyces coelicalor] Steptomyces coelicalor] Steptomyces coelicalor] Steptomyces coelicalor] Steptomyces coelicalor] 80071744 (2221, 2222) Novel Protein sim. GBank gij2622039 (AE000868) - type 1 restriction modification system, subunit S Methanobacterium thermoaulotrophicum] 82456352 (2225, 2224) Movel Protein sim. GBank gij083428 pitative oxidoreductase [Streptomyces coelicalor] 141023862) putative oxidoreductase [Streptomyces coelicalor] Novel Protein sim. GBank gij083428 pitative oxidoreductase [Streptomyces coelicalor] Novel Protein sim. GBank gij083428 pitative oxidoreductase [Streptomyces coelicalor] 141023862) putative oxidoreductase [Streptomyces coelicalor] 14103862) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286239 (2229, 2230) 1410286			\rightarrow			
1793 1410 (2211, 2212) 1793 1410 (2211, 2212) 1793 1410 (2211, 2214) 1794 21 Isple A34 16 ISECY_STRSC - PREPROTEIN 1794 21 Isple A34 16 ISECY_STRSC - PREPROTEIN 1794 21 Isple A34 16 ISECY_SUBUNIT 1784 NSL OCASE_SECY_SUBUNIT 1785 Novel Protein sim. GBank gil 222039 (AE000868) - type I 1785 Novel Protein sim. GBank gil 222039 (AE000868) - type I 1785 Novel Protein sim. GBank gil 22203 SIST Novel Protein sim. GBank gil 22183 Totel Novel Protein sim. GBank gil 23183 Totel Novel Protein sim. GBank gil 2485 Novel Protein	1105					264566
80470019 (2213, 2214) Novel Protein sim. GBank 80440616 (2215, 2216) Novel Protein sim. GBank 80140616 (2215, 2216) Novel Protein sim. GBank gil2995310 emb CAA18338 -	1106				UNCLASSIFIED	264555, 264369
80440616 (2215, 2216) Novel Protein sim. GBank gil2995310jemblCAA18338j - TRANSC - PREPROTEIN TRANSC CCASE SECY SUBUNIT TRANSC SCASION TRANSC SECY SUBUNIT (AL 022268) putative ATP-dependent helicase [Streptomyces coelicolor] Subunit S Subunit S Testriction modification system, subunit S Testriction modification system, subunit S Testriction modification modification sim. GBank gil2218376jemblCAA19628j - Secondary Subunit S Sub	1107					264906, 264769
gi 1173421 sp P43416 SECY_STRSC - PREPROTEIN	1108		Novel Protein sim. GBank		transport	264907 264510 264511 264600 264602
80064615 (2217, 2218) Novel Protein sim. GBank gil2995310 emb CAA18338 - (AL022268) putative ATP-dependent helicase [Streptomyces coelicalor] 80503554 (2219, 2220) Restriction modification system, subunit S [Methanobacterium thermoautotrophicum] 95010088 (2223, 2224) R2456352 (2228, 2226) Novel Protein sim. GBank gil3218376 emb CAA19628 - R44023862) putative oxidoreductase [Streptomyces coelicolor] 14998014 (2227, 2228) Novel Protein sim. GBank gil1083428 pir [S54876 - NAD(P) + transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse 1765583 (2229, 2230) 178841152 (2231, 2222)			gil173421jsp P43416 SECY_STRSC-PREPROTEIN TRANSLOCASE SECY SUBUNIT			264605, 264768, 264769
Streptomyces coelicolor Streptomyces coelicolor	109	80064615 (2217, 2218)	Novel Protein sim. GBank gij2995310 emb CAA18338 - (AL022268) putative ATP-dependent helicase		helicase	264602, 264605, 264636
80503554 (2219, 2220) 806071744 (2221, 2222) Novel Protein sim. GBank gi[2622039 (AE000868) - type I restriction modification system, subunit S [Methanobacterium thermoautotrophicum] 82456352 (2225, 2224) 82456352 (2225, 2226) Novel Protein sim. GBank gi[3218376 emb CAA19628 - UNCLASSIFIED (AL023862) putative oxidoreductase [Streptomyces coelicolor] 14998014 (2227, 2228) Novel Protein sim. GBank gi[1083428 pir [1554876 - NAD(P)+ transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse 11765583 (2229, 2230) 79841152 (2231, 2232)						
80071744 (2221, 2222) Novel Protein sim. GBank gi[2622039 (AE000868) - type 1 restriction modification system, subunit S [Methanobacterium thermoautotrophicum] 82456352 (2225, 2226) Novel Protein sim. GBank gi[3218376 emb CAA19628 - (AL023862) putative oxidoreductase [Streptomyces coelicolor] 14998014 (2227, 2228) Novel Protein sim. GBank gi[1083428 pir [1554876 - NAD(P) + transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse 11765583 (2229, 2230) 79841152 (2231, 2232)	<u>=</u>					264908, 264593, 265010, 264601, 264603,
80071744 (2221, 2222) Novel Protein sim. GBank gil2622039 (AE000868) - type I restriction modification system; subunit S [Methanobacterium thermoautotrophicum] 95010088 (2223, 2224) 82456352 (2225, 2226) Novel Protein sim. GBank gil3218378 emb CA419628 - UNCLASSIFIED (AL023862) putative oxidoreductase [Streptomyces coelicolor] 14998014 (2227, 2228) Novel Protein sim. GBank gil1083428 pir S54876 - NAD(P)+ transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse (UNCLASSIFIED (UNCLASSIFIED (2231, 2232))						264604, 264605, 264682, 264769, 264693, 264636
Methanobacterium thermoautotrophicum 95010088 (2223, 2224)	1111					18108370, 264557
95010088 (2223, 2224) 82456352 (2225, 2226) Novel Protein sim. GBank gil3218376 emb CAA19628 - (AL023862) putative oxidoreductase [Streptomyces coelicolor] 14998014 (2227, 2228) Novel Protein sim. GBank gil1083428 pir S54876 - NAD(P) * transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse transhydrogenase (B-specific) (EC 2.6.1.1) 1765583 (2229, 2230) 178841152 (2231, 2232)			restriction modification system, subunit S [Methanobacterium thermoautotrophicum]			
82456352 (2225, 2226) Novel Protein sim. GBank gij3218376 emb CAA19628 - (AL023862) putative oxidoreductase [Streptomyces coelicolor] 14998014 (2227, 2228) Novel Protein sim. GBank gij1083428 pir S54876 - NAD(P)* transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse 11765583 (2229, 2230) UNCLASSIFIED 178841152 (2231, 2232)	1112					264908
(AL023862) putative oxidoreductase [Streptomyces coelicolor] 14998014 (2227, 2228) Novel Protein sim. GBank gil1083428 pir S54876 - NAD(P)+ transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse	1113				UNCLASSIFIED	264600, 264602, 264604, 264605, 264762.
14998014 (2227, 2228) Novel Protein sim. GBank gil1083428 pir SS4876 - NAD(P)+ transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse 11765583 (2229, 2230) 79841152 (2231, 2232)			(AL023862) putative oxidoreductase [Streptomyces coelicolor]			264769, 264565
NAD(P)+ transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse UNCLASSIFIED	1114		Novel Protein sim. GBank gij1083428 pir S54876 -		dehydrogenase	264636
precursor - mouse UNCLASSIFIED			NAD(P)+ transhydrogenase (B-specific) (EC 1.6.1.1)			
UNCLASSIFIED			_			
	1115	11765583 (2229, 2230)			UNCLASSIFIED	264686
	1116	79841152 (2231, 2232)				264908

18108392, 264488, 263994, 264489, 18108398, 56182575, 22278995, 22278996, 35696288, 56994075, 22278997, 22278998, 22278999, 224259, 29331822, 56182181, 29331824, 66714117, 29331828, 3569652, 33656970, 2614649, 264508, 264508, 264509, 2564006, 265007, 264510, 264509, 52644045, 56182435, 264510, 264509, 526410, 60170831, 264591, 264592, 2655009, 264910, 60170831, 264591, 264592, 264593, 60433356, 265594, 60433438, 264593, 60433356, 264596, 264759, 264504, 23109954, 33657064, 265011, 87188559, 264600, 264601, 264602, 264603, 265017, 264604, 265018, 264762, 264766, 264760, 55811150, 264681, 264768, 264766, 264760, 55811150, 264681, 264768, 264766, 264760, 5581180, 264681, 264768, 264766, 264760, 26488, 264682, 264768, 264766, 264688, 264682, 264682, 264768, 264766, 264888, 264682, 264682, 264768, 264766, 264681, 18108356, 284689, 264691, 265021, 265022, 60170615, 264691, 33657023, 284692, 284689, 6274620, 33657023, 284692, 284689, 6274620, 33657023, 284692, 284689, 6274620, 33657023, 284692, 284689, 6274620, 33657023, 284692, 284689, 6274620,					264693, 27486265 264693, 27486265	0.264909			ĺ	264511
helicase	UNCI ASSIFIED	UNCLASSIFIED	collagen	UNCLASSIFIE	UNCLASSIFIED	INCI ASSIFIED	UNCLASSIFIED	hqa	UNCLASSIFIED	polymerase
Contains protein domain (PF00270) - helicase DEAD/DEAH box helicase				Contains protein domain (PF00013) - UNCLASSIFIED KH domain						Contains protein domain (PF00476) - polymerase DNA polymerase family A
1117 95305465 (2233, 2234) Novel Protein sim. GBank gil3255965 emb CAA94089 - (Z70200) U5 snRNP-specific 200KD protein [Homo sapiens] DEAD/DEAH box helicase			79480463 (2239, 2240) Novel Protein sim. GBank gil5420387 emb CAB46679.1 - [(AJ243459) proteophosphoglycan [Leishmania major]	79471716 (2241, 2242) Novel Protein sim. GBank gi 1644450 (U67864) - MEX-3 Caenorhabditis elegans	79637119 (2245, 2246) Novel Protein sim. GBank gil98800 pir. S17768 - 3-dehydroquinate synthase (EC 4.6.1.3) - Mycobacterium tuberculosis			79758914 (2251, 2252) Novel Protein sim. GBank gij138154 sp P03643 VGG_BPPHX - MAJOR SPIKE PROTEIN (G PROTEIN) (GPG)		Novel Protein sim. GBank gi 5002704 emb CAB4358.1 - (AJ242630) DNA polymerase [Methylobacterium sp. DM4]
95305465 (2233, 2234)	Į	_			79637119 (2245, 2246)	79811596 (2247, 2248)	1125 79757861 (2249, 2250)	79758914 (2251, 2252)	11800930 (2253, 2254)	8364885 (2255, 2256)
	1118	1119	1120	1121	1123	1124	1125	1126	1127	1128

	80422480 (2257, 2258)	80422480 (2257, 2258) Novel Protein sim. GBank gil5689485 dbj BAA83026.1 - (AB028997) KIAA1074 protein [Homo sapiens]	Contains protein domain (PF00170) - UNCLASSIFIED bZIP transcription factor		265011, 264766
1130	79420151 (2259, 2260)				264595
	80055391 (2261, 2262)	80055391 (2261, 2262) Novel Protein sim. GBank gil4981328 gb AAD35881.1 AE00174 - (AE001747) bioY protein [Thermotoga maritima]		UNCLASSIFIED	35696286, 22278998, 29331828, 264603, 264605, 264559
1132	82062248 (2263, 2264)	82062248 (2263, 2264) Novel Protein sim. GBank gij1841552 (U89336) - unknown [Homo sapiens]		UNCLASSIFIED	22278996, 264906, 265009, 264600, 264602, 264604, 264605, 264760, 32833986, 18108374
1133	17290437 (2265, 2266)			UNCLASSIFIED	265018
1134	_			UNCLASSIFIED	264512, 264534
1135		80029393 (2269, 2270) Novel Protein sim. GBank gil4539171(emb CAB39700.1) - (AL049485) conserved hypothetical protein [Streptomyces coelicolor]			264508, 264600, 264602, 264603, 13108376
1136	79842052 (2271, 2272)	79842052 (2271, 2272) Novel Protein sim. GBank gil4982454[gb AAD36931.1 AE00182 - (AE001823) ATP- dependent protease LA, putative [Thermotoga maritima]		UNCLASSIFIED	264906, 264908
1137	90931557 (2273, 2274)	90931557 (2273, 2274) Novel Protein sim. GBank gil4972746[gb]AAD34768.1 -	Contains protein domain (PF00515) - collagen	collagen	22278998, 22278999, 35696052, 264907,
		(A⊦132160) unknown [Drosophila melanogaster]	TPR Domain		265009, 6043336, 264596, 265010, 264448, 264682, 264682, 264692, 264689, 265020, 264692,
					558115/6, 35695855, 264631, 264532, 22279002
1138	79841163 (2275, 2276)	79841163 (2275, 2276) Novel Protein sim. GBank		struct	264908
		gij731607ispjP38739 YHC8_YEAST - HYPOTHETICAL 63.8 KD PROTEIN IN GUT1-RIM1 INTERGENIC REGION PRECURSOR			
1139	79633561 (2277, 2278)	79633561 (2277, 2278) Novel Protein sim. GBank gi 3650031 (AC005396) - putative protein [Arabidopsis thaliana]		UNCLASSIFIED	264693
1140	39480358 (2279, 2280)			UNCLASSIFIED	264593
1141	79638019 (2281, 2282)			UNCLASSIFIED	265019, 264693
1142	19635848 (2283, 2284)			UNCLASSIFIED	264631
1143		87762158 (2285, 2286) Novel Protein sim. GBank gil3928000 emb CAA05880 - [(AJ003125) procollagen I N-proteinase [Homo sapiens]	Contains protein domain (PF00090) - oxidase Thrombospondin type 1 domain	oxidase	56182575, 264908, 264600, 264632, 87168518
1144	80088988 (2287, 2288)			UNCLASSIFIED	264635, 264636, 264907, 264593, 264908, 264566, 264909
1145	14610262 (2289, 2290)			UNCLASSIFIED	264112
1146	82062092 (2291, 2292)			UNCLASSIFIED	264769, 264689, 35696286, 264760, 264905, 264486, 264559
1147	80071761 (2293, 2294)				264557
1148	80048433 (2295, 2296)	80048433 (2295, 2296) Novel Protein sim. GBank gi[2499003 sp P76422 THID_ECOLI - PHOSPHOMETHYLPYRIMIDINE KINASE (HMP- PHOSPHATE KINASE) (HMP-P KINASE)		kinase	264591
1149		11607438 (2297, 2298) Novel Protein sim. GBank gij2896734 emb CAA17213.1 - (AL021897) hypothetical protein Rv1097c [Mycobacterium tuberculosis]			264591

150	81325074 (2299, 2300)	1150 [81325074 (2299, 2300) [Novel Protein sim. GBank gil2895095 (AF011337) - putative		ATPase_associated	ATPase_associated 264488, 35696286, 264907, 264908, 264909,
		E1-E2 ATPase [Mus musculus]			264910, 264593, 264596, 264758, 264764.
					264766, 264768, 264693, 264628, 60431850.
					264564, 264566, 264567
1151	80070874 (2301, 2302) Novel Protein sim. Gl	Novel Protein sim. GBank gil4324655 gb AAD16978 -		polymerase	264595
		(AF108191) DNA polymerase III alpha subunit			
		[Streptomyces coelicator]			
1152	80235547 (2303, 2304)	80235547 (2303, 2304) Novel Protein sim. GBank gij3874275jembjCAB07311.1j -		glycoprotein	264488, 22278998, 264905, 264629, 264486
		(Z92825) predicted using Genefinder; Similarity to Yeast low			
	-	afinity glucose transporter HXT4 (PS:32467); cDNA EST			
		EMBL:C12555 comes from this gene; cDNA EST			
		lyk404c10.3 comes from this gene; cDNA EST yk404c10.5			
		comes from thi			
1153	80027783 (2305, 2306)	-			264910, 264555, 264557
1154	83002995 (2307, 2308) Novel Protein sim. G	Novel Protein sim. GBank gil4240315 dbj BAA74936.1 -		UNCLASSIFIED	265008
1		(AB020720) KIAA0913 protein [Homo sapiens]			
25	79411098 (2309, 2310)	1155 79411098 (2309, 2310)		UNCLASSIFIED	264690, 264636
ဖြွ	57147843 (2311, 2312)		Contains protein domain (PF00122) - transport	transport	264603
		N_ECOLI - ZINC- E (ZN(II)-TRANSLOCATING P-	E1-E2 ATPase		
		TYPE ATPASE)			
1157	95287711 (2313, 2314)	95287711 (2313, 2314) Novel Protein sim. GBank		UNCLASSIFIED	264906, 264907, 264758, 264766, 264769,
		gi 418480 sp P32139 YIHR_ECOLI - HYPOTHETICAL 34.0			264689, 264638, 264566
		KD PROTEIN IN GLNA-RBN INTERGENIC REGION			
1158	82454917 (2315, 2316)	82454917 (2315, 2316) Novel Protein sim. GBank			264906, 264762, 264687, 264769, 264689.
		gi 2496481 sp Q50724 Y09S_MYCTU - HYPOTHETICAL 87.3 KD PROTEIN CY78.27C			18108374, 35695855
1159	79186451 (2317, 2318)				264687
1160		01220803 (2310 2320) Movel Protein eim GRank mitt 13640 Richil Ba A 114901.			29331822, 29331824, 66714117, 29331826.
3		(D79995) similar to pig tubulin-tyrosine ligase. [Homo			60433356, 265018, 265019, 83373044,
		sapiens]			18108385, 22279000, 22279002, 264563
1161					264369
25	_	79635357 (2323, 2324) Novel Protein sim. GBank gi 2443342 dbj BAA22380 - (D88764) alpha 2 type I collagen [Rana catesbeiana]		UNCLASSIFIED	264693
1163	79563186 (2325, 2326)	79563186 (2325, 2326) Novel Protein sim. GBank gil4503375 reflNP 001376.1 pDPYS - dihydropyrimidinase		UNCLASSIFIED	29331827, 264906
20	_	Novel Protein sim. GBank		transport	18108398, 29331827, 29331828, 29146498,
		gij5052554lgb AAD38607.1 AF14563 - (AF145632) BcDNA GH06032 [Drosophija melanogaster]			29146499, 18108354, 21906768, 29148627, 21906769, 264693, 18108382, 18108385
1165	80491888 (2329, 2330)	+	The state of the s	UNCLASSIFIED	264602, 264605, 264769, 18108370, 18108374, 264565
1166		88096456 (2331, 2332) Novel Protein sim. GBank gil4589476 dbj BAA76766 1 -			264488, 35696286, 22278999, 264259.
		(AB023139) KIAA0922 protein [Homo sapiens]			66714117, 60432289, 35696052, 264905,
					56182435, 265006, 60433438, 264759,
					21906754, 33109954, 265017, 265019.
					264448, 264288, 264766, 264685, 35696423,
					35695855, 264558, 18108385, 60432113

1167		79963862 (2333, 2334) Novel Protein sim. GBank gilz580433ldbjjBAA23138j - IO76414) mGm hydrolase (Stanbylococus aurous)		kinase	264488
1168	88094678 (2335, 2336)	4-			
				UNCLASSIFIED	264259, 29331827, 56182435, 60433438, 265010, 3265702, 26608888, 26466
1169				INCIASSIED	203019, 33037023, 33083033, 204300
1170	į			INCI ASSIEIED	264603
1171		Novel Protein sim. GBank gil2772914 (AF029249) -		INCLASSIFIED	204602
ļ	_	precollagen D [Myti			204330
1172		Novel Protein sim. GBank		UNCLASSIFIED	55810764, 35696052, 264634, 264486
		gi 4757846 ref NP_004317.1 pBCL9 - B-cell CLL/lymphoma 9			
1173	80235713 (2345, 2346) Novel Protein sim.	Novel Protein sim. GBank gil2564053ldbilBAA229461.			100100 000100 E00100 000100 000100
]	\neg	- Æ I			zbabus, zbasus, zbasuv, zbasus, zbasel, 264632, 264638, 264639
11/4	20293077 (2347, 2348) Novel Protein sim. (Novel Protein sim. GBank gi 2911027 emb CAA17520 -		dehydrogenase	264600
1476		(AL021958) mmsA (Mycobacterium tuberculosis)		•	
2		Novel Protein Sim. GBank	Contains protein domain (PF00205) - carboxylase	carboxylase	264601
_		July 183333 DCSZ34 DCRZ = NDOLE-3- PYRUVATE DECARBOXYLASF (INDOLEPYR) INATE	Thiamine pyrophosphate enzymes		
		DECARBOXYLASE)			
1176		. (90	Contains protein domain (PE00586) - synthase	synthasa	264500 264005 264503 264602 264605
			AIR exothese related exptain	Depillinée.	zu-sus, zo-sus, zo-sus, zo-touz, zo-tous
		sis]	Any symmase related protein		
1177	80064647 (2353, 2354)		Contains a second of the contact of	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
		G CUPLA - 3-OXOACYL-IACYL-	Contains protein domain (PEDU105) - reductase short chain dehydrogenase	reductase	264605
	_	KETOACYL-ACYL CARRIER PROTEIN REDUCTASE)			
<u> </u>			Contains protein domain (PF00122) -	ATPase_associated	Contains protein domain (PF00122) - ATPase_associated 65274572, 18108398, 22278998, 22278999,
		gipos roar (regine uposes, riphica - raminal intrahepatic	E1-E2 ATPase		29331826, 264508, 264908, 264828,
		cholestasis 1, (progressive, Byler disease and benign			33657402, 33109954, 264769, 21906765
		recurrent)			21906766, 21906768, 55811957, 33657023,
_					264629, 55811576, 35696423, 264636,
					264556, 56182323, 60432113, 22279000,
1179		17988.11 -	Contains protein domain (PE00496) - Iransport	fransnort	264603
	45		Bacterial extracellular solute-binding		
1180	11794446 (2359, 2360)	11794446 (2359, 2360) Novel Protein sim. GBank gil2558614 lemblC A A A A 7871	Contains profession domain (DEC) 1920		
	•		Condins protein domain (PFU1220) - synthase	synthase	264638
			Denydroquinase class II		
1181	17946362 (2361, 2362)				- 70100
1182	81494264 (2363, 2364)	81494264 (2363, 2364) Novel Protein sim, GRank ail Ed 201387 Jomes I Cabace 2011		UNCLASSIFIED	7,000
		(AJ243459) proteophosphoglycan [Leishmania major]			265007, 265009, 264564, 264909, 264693
183	79574044 (2365, 2366)				264680 35606422 264620 10400205
	52559933 (2367, 2368)	52559933 (2367, 2368) Novel Protein sim. GBank gil4091877 (AF061331) - alpha		INC. ACCIETED	264603
$\overline{}$		galactosidase precursor (Saccharopolyspora erythraea)			204002
1185	79491185 (2369, 2370)	Novel Protein sim. GBank gi[2129478 pir 551939 -		glycoprotein	263967
		chimase (EC 3.2.1.14) precursor - beet			

1186 20224012 (2371, 2372) 1187 79248834 (2373, 2374) 1188 79831387 (2375, 2375) Novel Protein sim Chank ditrockorao Accounces	1) Nove Protein sim GBank ull 2006030 (ACDE	675)		UNCLASSIFIED UNCLASSIFIED	264559 29331825, 265017, 18108351
	hypothetical protein (Synechococcus PCC7002)	- ()		UNCLASSIFIED	264905, 264906
					264692
1191 80310105 (2381, 2382)				UNCLASSIFIED	265018
1100 12504644 10300 0004				UNCLASSIFIED	264600, 264605, 264764, 35695855, 264638, 264486
			Contains protein domain (PF01581) -		264636
				UNCLASSIFIED	263978
	Novel Protein sim. GBank gi 854065 emb CA/4 (X83413) U88 (Human herpesvirus 6)	58337 -		UNCLASSIFIED	265007, 265008
_	Novel Protein sim. GBank gil 1790277 (AE0004 oxidoreductase [Escherichia coli]	59) - putative		UNCLASSIFIED	264605
1195 13000688 (2391, 2392)					264689
gilz497360 Movel Frotein sim. GBank gilz497360 sp QS0715 IMDH_MYCTU - INOSINE-5: MONOPHOSPHATE DEHYDROGENASE (IMP DEHYDROGENASE) (IMPDH) (IMPD)	JINOVEL FYOREIN SIM., CHEMINA GI[2497360]SPIGS0715]IMDH_MYCTU - INOSINI MONOPHOSPHATE DEHYDROGENASE (IMP DEHYDROGENASP) (IMPDH) (IMPD)		Contains protein domain (PF00571) - dehydrogenase CBS domain	dehydrogenase	264594
95290101 (2395, 2396)					264603
1199 (81882011 (2397, 2398) Novel Protein sim. GBank	Novel Protein sim. GBank				20000
gil709525[sp P54673]P3K1_DICDI - PHOSPHATIDYLINOSITOL 3-KINASE 1 (P13-KINASE) (PTDINS-3-KINASE) (P13K)	gij1709525[sp P54673]P3K1_DICDI - PHOSPHATIDYLINOSITOL 3-KINASE 1 (P13-K (PTDINS-3-KINASE) (P13K)	INASE)			264259, 264757, 33109954, 21906768
1200 3040000 (2359, 2400)				UNCLASSIFIED	264910
	Novel Protein sim. GBank gil2499877IspIP70645 BLMH_RAT - BLEOMY HYDROLASE (BLM HYDROLASE) (BMH)	N.		cathepsin	264766, 264769
	nover Protein sim. GBank gil606342 (U18997) reading frame open far upstream of start; poss frameshift, linking to previous ORF [Escherichi	- ORF_0622; lible a coli]		ribosomalprot	264600, 264558
821253/3 (2405, 2406)				UNCLASSIFIED	264768, 264769, 35695917, 264910, 264760,
1204 80503016 (2407 2400) No. 15	O CONTRACTOR				zo4906, zo4907, zb46z9, z64908, z64909, 264766
	NOVER PROBEITS STANK gil2500728 sp Q59912 SECY_STRGB - PREPRC TRANSLOCASE SECY SUBUNIT	OTEIN			264905, 264769, 264636
				UNCLASSIFIED	264566
					264556 264557 264558
79841192 (2413, 2414)					29331824 264909 265021 18108370
or rooz ir (z4 t5, z4 t5) Inovel Protein sim. GBank gil2645560 (AF027954) - Bci-2- related ovarian killer protein [Rattus norvegicus]	Novel Protein sim. GBank gil 2645560 (AF027 related ovarian killer protein [Rattus norvegio:		Contains protein domain (PF00452) - apoptosis Apoptosis regulator proteins, Bcl-2		29331824, 29331825, 29331827, 265007, 264764, 264683, 264769, 264688, 264689
			tarnuy		

6021	1209 (9185/42 (2417, 2418) Novel Protein sim. G	gilt175033lsp P44398 XYLA_HAEIN - XYLOSE	Contains protein domain (PF00259) - isomerase Xylose isomerase	isomerase	264687, 264688
,		ISOMERASE			
1210	56426884 (2419, 2420)			UNCLASSIFIED	264907, 264693
1211		94665655 (2421, 2422) Novel Protein sim. GBank giļ421095[pir]S30688 - hypothetical protein o246 - Escherichia coli		transferase	264591, 264592, 264595
1212	-		Contains protoin domain (DE01113)		DEACON DESCRIP
4		(293785) predicted using Genefinder; similar to RNA	Putative GTP-ase activating protein		204069, 203907
		recognition motif. (aka RRM, RBD, or RNP domain); cDNA	for Art		
		EST EMBL:T01682 comes from this gene; cDNA EST			
		EMBL:M75823 comes from this gene; cDNA EST			
		EMBL:D27559 comes from this ge			
1213	_	79859633 (2425, 2426) Novel Protein sim. GBank gill26292Inrfl11505375A - vir		kinase	264909
!		gene (Bordetella pertussis)			20203
1214	-	10144306 (2427 2428) Novel Protein sim GBank		Calabo ONI I	364008
		gi5726285[gb]AAD48396,1]AF12616 - (AF126162) HERV-H			
		LTR associating protein 2 [Homo sapiens]			
1215	80050106 (2429, 2430)			UNCLASSIFIED	265009, 264601, 264602, 264603, 33657109
	_	(Z98268) recN [Mycobacterium tuberculosis]			
1216	_	20438324 (2431, 2432) Novel Protein sim. GBank		transferase	264604
		dil417329lsplP33038lMURA_ENTCL - UDP-N-			
		ACETYLGLUCOSAMINE 1-			
		CARBOXYVINYI TRANSFERASE (FNOVI PYBLIVATE			
		TRANSFERASE (100-N-ACETY) COLLOCSAMINE			
		ENOLPYRUVYL TRANSFERASE) (FPT)			
1217	95011344 (2433, 2434)	+		INCLASSIED	26490F 264907 264908 264501 264766
				7111000000	201900, 201901, 201900, 201991, 201700, 264601 264601 264620 264630 264636
					264564
1218	Į	11093680 (2435, 2436) Novel Protein sim. GBank qi11805460idbilBAA09022I -		dehydrogenase	264601
		(D50453) homologue of succinate semialdehyde dehydrogenase GabD of E. coil (Bacillus subtilis)			
1219	91216252 (2437, 2438)	-			56181686, 29331822, 60432289, 264601,
	_				264692, 264629
1220		91241524 (2439, 2440) Novel Protein sim. GBank gil4240315 dbj BAA74936.1 -		oncogene	52644507, 264905, 264909, 265008, 265019,
		(AB020720) KIAA0913 protein [Homo sapiens]			265020, 52644150, 33657023, 264693,
					3303/182, 33093/03, 204634, 222/9000, 22279002, 264482
1221	183045055 (2441 2442) Novel Protein sim G	Novel Protein sim GBank dil 2143888 hirti 152523		INCLASSIED	264768 265020 264006
		Incleoporin p62 homolog - rat (fragment)		Osicios Company	204/00, 203020, 204300
1222	į	20711865 (2443, 2444) Novel Protein sim. GBank	Contains protein domain (PF00486) - phosphatase	phosphatase	264601
		gi[730805]spiP39663JSPHR_SYNP7 - ALKALINE	Transcriptional regulatory protein, C	•	
		PHOSPHATASE SYNTHESIS TRANSCRIPTIONAL	terminal		
1223	11615647 (2445, 2446)				264593
1224	80432645 (2447, 2448)	1224 80432645 (2447, 2448) Novel Protein sim. GBank	Contains protein domain (PF01472) -	kinase	264593 264600 264601 264603 264605
		gil172627[splP46546]PROB_CORGL - GLUTAMATE 5- KINASE (GAMMA-GI ITAMY) KINASE (GAMMA-GI ITAMY)	PUA domain		264768, 18108376, 264635, 18108387

1225	180434427 (2449 2450)					
1228	1226 80237518 (2451 2452)	Name Brates also OB House Com.			264768	_
	(5431, 5435,	Cocoro (2701, 2702) Nover Floren sin. Gbank gilz 105050[emb]CAB08836 - (2701, 2701, 2702) Nover Floren Sin. Gbank gilz 105050 [Mycobacterium floren Rv3644c	-	polymerase	264905, 264512, 264689	
1227		78422138 (2453, 2454) Novel Protein sim. GBank gil 706768jspp98133jFBN1_BOVIN - FIBRILLIN 1 PRECI IRSOR AMPANO		UNCLASSIFIED	264908, 264637, 264639	
1228		79209027 (2455, 2456) Novel Protein sim. GBank gil1653901 dbj BAA18811 - (D90917) acriflavine resistance protein [Synechocystis sp.)	Contains protein domain (PF00873) - AcrB/AcrD/AcrF family		264605, 264634	
1229			1	UNCLASSIFIED	87168474, 265011, 87168559, 264681, 284889, 264693, 65274630, 49409374	· · · ·
1230		80049357 (2459, 2460) Novel Protein sim. GBank gil116230 sp P28598 CH60_BACSU - 60 KD CHAPERONIN TCP-1/cpn60 chaperonin family (PROTEIN CPNA) (GROEI PROTEIN)	Contains protein domain (PF00118) - eph TCP-1/cpn60 chaperonin family	eph	264909, 264605, 18108388	
1231	\mathbf{T}			INCI ASSIFIED	264008	
1232	_	79853104 (2463, 2464) Novel Protein sim. GBank gij1215733 (U48718) - OphC [Agrobacterium tumefaciens]		transport	264909	
1233		80255179 (2465, 2466) Novel Protein sim. GBank gilf16298 sp p20730 CHHC_BOMMO - CHORION CLASS HIGH-CYSTEINE HCB PROTEIN 13 PRECURSOR (HC- B.13)		UNCLASSIFIED	265017, 264564	
1234		Novel Protein sim. GBank gi 729671 sp P40280 H2A_MAIZE - HISTONE H2A	Contains protein domain (PF00125) - histone Core histone H2A/H2B/H3/H4	histone	265008, 265010, 18108381	
1230	_				264634, 264762	_
250				UNCLASSIFIED	265018, 55811150, 264565, 264757	_
25		833/1/82 (24/3, 24/4) Novel Protein sim. GBank gi]3875133]emb[CAA94750] - (Z70750) similar to actin binding domain; cDNA EST EMBL.T00093 comes from this gene; cDNA EST EMBL.D34443 comes from this gene; cDNA EST EMBL.D37508 comes from this gene; cDNA EST EMBL.D37508 comes from this gene; cDNA EST EMBL.D447 comes from this gene; cDNA EST EMBL.D64247 comes from this gene; cDNA EST EMBL.D			264755, 264559 264555, 264559	
1238		Novel Protein sim. GBank gij3895470 (AF061443) - G protein-coupled receptor LGR4 [Rattus norvegicus]	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat	glycoprotein	264259, 29331822, 29331824, 35696052, 264508, 264906, 52644045, 52646317, 264288, 264769, 264693, 264632, 264634,	
1239		82197449 (2477, 2478) Novel Protein sim. GBank gil4007990 gb AAC95339 - (AF084363) DOK protein [Mus musculus]		oncogene	264509, 264511, 264759, 264760, 264764, 264857	
1240	80497259 (2479, 2480)	80497259 (2479, 2480) Novel Protein sim. GBank gi[176192[sp]P45420]YHCD_ECOLI - HYPOTHETICAL OUTER MEMBRANE USHER PROTEIN IN GLTF-NANT INTERGENIC REGION PRECURSOR			264769	
1241		80020711 (2481, 2482) Novel Protein sim. GBank gi r21383 sp p19904 GLNA_VIBAL - GLUTAMINE SYNTHETASE (GLUTAMATE-AMMONIA LIGASE)	Contains protein domain (PF00120) - UNCLASSIFIED Glutamine synthetase	UNCLASSIFIED	264601, 264604, 264638	
1242	1242 79775890 (2483, 2484)				264906, 264907, 264908, 264634	

1243	1243 79779458 (2485, 2486) Novel Protein sim. G	Novel Protein sim. GBank gi[3355671 emb CAA19971 -		UNCLASSIFIED	18108374 35695917 35695855 265009
		(AL031124) branched-chain amino acid aminotransferase			264508, 264909
		[Streptomyces coelicolor]			
1244	10284821 (2487, 2488)	10284821 (2487, 2488)}Novel Protein sim. GBank gij2970646 (AF051945) - Xin [Mire mineculus]		UNCLASSIFIED	264691
1245	80437103 (2489 2490)	80437103 (2489, 2490) Novel Protein sim CB2nt vild58633814hilB&A76357 11	Cathor protocol distance		001700
2	(2403, 2430)	(AB016787) cytochrome o ubiquinol oxidase B	Contains protein domain (PF00115) - oxidase Cytochrome C and Quinol oxidase	oxidase	264768
		[Pseudomonas putida]	polypeptide I		
1246	80059321 (2491, 2492)	Novel Protein sim. G		UNCLASSIFIED	264604, 264636, 264557, 264564
		(ALUSTO41) purative pnenylalanyi-tikina synthetase beta ichain [Streptomyces coelicolor]			
1247	80064831 (2493, 2494)	80064831 (2493, 2494) Novel Protein sim. GBank gi[2621684 (AE000842) -			264758 264605 264639
		adhesion protein [Methanobacterium thermoautotrophicum]			
1248	88070353 (2495, 2496)	88070353 (2495, 2496) Novel Protein sim. GBank	Contains protein domain (PF00316) - UNCLASSIFIED	UNCLASSIFIED	18108392, 264259, 29331826, 264106,
		gi 1352403 sp P09467 F16P_HUMAN - FRUCTOSE-1,6-	Fructose-1-6-bisphosphatase		264508, 264907, 264828, 265009, 60433356,
		BISPHOSPHATASE (D-FRUCTOSE-1,6-BISPHOSPHATE			264757, 264758, 21906754, 265010, 265011,
					265018, 265019, 264760, 18108351, 18108354, 265021, 18108376, 18108377.
1					264630, 18108385
1249	80056657 (2497, 2498)	_		transport	264908, 265010, 264600, 264603, 264691,
		(AL021184) hypothetical protein Rv1473 [Mycobacterium			18108376
		tuberculosis			
1250		12694385 (2499, 2500) Novel Protein sim. GBank		UNCLASSIFIED	264689
		gillz/85jspjP05100j3MG1_ECOLI - DNA-3-			
		METHYLADENINE GLYCOSIDASE I (3-METHYLADENINE-			
		DNA GLYCOSYLASE I, CONSTITUTIVE) (TAG I)			
1251	79850448 (2501, 2502)				264000
1252	79458087 (2503 2504)			100 to 10	201200
707	1945001 (2303, 2304)			UNCLASSIFIED	264683, 263976
6671	80050121 (2505, 2506)	lovel Protein sim. C		glycoprotein	264600, 264603, 18108376
		gij56/01/6lgb AAD46616.1 AF16131 - (AF161317) NRAMP			
		manganese transport protein mitra (Salmonella typhimurium)			
1254	87716767 (2507, 2508)	87716767 (2507, 2508) Novel Protein sim. GBank gij103160 pir S22126 - finger		UNCLASSIFIED	35696286, 264910, 264764, 264688.
		protein unkempt - fruit fly (Drosophila melanogaster)			21906767, 55811957, 264692, 264556,
1255	79169728 (2509, 2510)				204039
1256	87889508 (2511, 2512)	87889508 (2511, 2512) Novel Protein sim. GBank ail 2995353 lembl CAA04608 11-		INCI ASSIFIED	60432389 264600 264605 264764 264687
		(AJ001206) pep2 [Streptomyces coelicolor]			264769, 264689, 27486265, 18108374,
					18108376
1257	80201435 (2513, 2514)	80201435 (2513, 2514) Novel Protein sim. GBank gij3193306 (AF069300) -		UNCLASSIFIED	264094, 265019
		Inducible-like protein (GB:AL021637) IArabidopsis thaliana			
1258	20708150 (2515, 2516)			UNCLASSIFIED	264602, 263978
1259	80186012 (2517, 2518)			UNCLASSIFIED	264906, 264448, 264908
1260	80084606 (2519, 2520)				264634 264639
					2000

1261	87412802 (2521, 2522) Novet Protein sim. G (AB029010) KIAA10	Novel Protein sim. GBank gi 5689511 dbi BAA83039.1 - (AB029010) KIAA1087 protein [Homo sapiens]	Contains protein domain (PF01699) - cadherin Sodium/calcium exchanger protein	cadherin	29331824, 264906, 264909, 264768, 264769, 264689, 264693, 264639, 18108384, 264563
1262		13504589 (2523, 2524) Novel Protein sim. GBank gil95100 pir S21334 - hypothetical protein 4 - Agrobacterium tumefaciens		UNCLASSIFIED	264634
1263					264602
1264		Novel Protein sim. GBank gil3550958 (AF004840) - CDO [Rattus norvegicus]		struct	264634
1265				UNCLASSIFIED	264563
1266	Ī			UNCLASSIFIED	264766, 264636, 264638, 264567
1267		80558918 (2533, 2534) Novel Protein sim. GBank gi 1085002 pir S55056 - mitochondrial carrier protein DIF-1 homolog - Caenorhabditis elegans	Contains protein domain (PF00153) - transport Mitochondrial carrier proteins	transport	264259, 21906754, 264369
1268		88178473 (2535, 2536) Novel Protein sim. GBank gi 4886445 emb CAB43370.1	Contains protein domain (PF00583) - UNCLASSIFIED Acetyltransferase (GNAT) family	UNCLASSIFIED	18108398, 22278995, 56994075, 60424269, 29331827, 264109, 264512, 265007, 265008, 265009, 264595, 33109954, 33657084, 87168559, 264600, 265018, 265019, 264369, 264688, 21906767, 265020, 52644150, 264691, 33657023, 33657349, 18108374, 264556, 18108385, 60432113, 22279002,
	_				264486
1269		79821946 (2537, 2538) Novel Protein sim. GBank gi 3334791 emb CAA19939 - (AL031107) hypothetical protein SC5A7.10c Streptomyces coelicolor]		UNCLASSIFIED	264508, 264905, 264906, 264667, 264693
1270		80031420 (2539, 2540) Novel Protein sim. GBank	Contains protein domain (PF01574) - dehydrogenase	dehydrogenase	265010, 264601
		gi 2851634 sp Q50591 Y0D1_MYCTU - HYPOTHETICAL 50.0 KD PROTEIN CY1A11.01	IMP dehydrogenase / GMP reductase N terminus		
1271				ATPase_associated 35696052, 264908	35696052, 264908
1272					264686, 264689
1273		80220315 (2545, 2546) Novel Protein sim. GBank gil1655665 emb CAB03731 -		UNCLASSIFIED	264509, 264639
		(281368) hypothetical prolein Rv2395 [Mycobacterium tuberculosis]			
1274				UNCLASSIFIED	264905, 264908, 264909, 264769
1275		20730763 (2549, 2550) Novel Protein sim. GBank	in domain (PF00183) -	eph	264602
	_	gij123726 spiP10413 HTPG_ECOLI - HEAT SHOCK PROTEIN HTPG (HIGH TEMPERATURE PROTEIN G) (HEAT SHOCK PROTEIN C62.5)	Hsp90 protein		
1276	_	21148644 (2551, 2552) Novel Protein sim. GBank gil2129478 pir S51939 - chilinase (EC 3.2.1.14) precursor - beet		UNCLASSIFIED	264369
1277				UNCLASSIFIED	264556
1278		11088365 (2555, 2556) Novel Protein sim. GBank		UNCLASSIFIED	264603
		giji i 5947 sjepji 449 posji AAJ_HAEIN - HYPO I HEI II CAL PROTEIN HI0183			
1279		21658756 (2557, 2558) Novel Protein sim. GBank gi 1929513 (U64318) - ATP synthase subunit beta (Moorella thermoacetical		synthase	264605
1280	79310959 (2559, 2560)			struct	263976

IFIED 29331825, 29331828, 264766, 83373044	FIED 265008	IFIED 264605		264604	264766, 264689, 263967	264605, 264639	IFIED 265011, 264602, 264766, 264687, 264769, 264689, 18108370, 264636, 18108385, 264563	35696052, 264107, 264508, 264509, 264905, 264905, 264906, 264907, 264908, 264909, 264510, 264911, 265007, 264512, 264910, 264511, 265007, 264595, 264758, 265011, 265019, 264760, 18108351, 264681, 264764, 264288, 264685, 264768, 264768, 264769, 265020, 265021, 264534, 264692, 18108370, 264628, 18108374, 35696423, 264555, 264555, 264555, 264555, 264555, 264555, 264555, 264555, 264555, 264555, 264555, 264555, 264555, 264556, 264557, 264558, 18108385, 264564,	264566, 264567, 264486, 18108391		265007, 265008, 264769	ctor 264092, 264259, 29331822, 29331824, 264508, 264906, 264909, 264512, 265008, 265009, 264361, 265019, 264369, 264388, 264686, 264639, 83373044, 22279002, 26482, 264563			nase 264689
- UNCLASS	UNCLASSIFIED	UNCLASSIFIED	-		- ebh	reductase	UNCLASSIFIED	- collagen	INC. ACCIEIED	- tm7	- helicase	- transcriptfa	UNCLASSIFIED	biotindep	dehydrogenase
Contains protein domain (PF00047) - UNCLASSIFIED	The model of the m				Contains protein domain (PF00183) - eph Hsp90 protein			Contains protein domain (PF00386) - collagen C1q domain		Contains protein domain (PF00001) - tm7 7 transmembrane receptor (rhodopsin family)	Contains protein domain (PF00271) - helicase Helicases conserved C-terminal domain	r Contains protein domain (PF01530) Zinc finger, C2HC type			
1281 94323988 (2561, 2562) Novel Protein sim. GBank gil1136501 (U39546) - surface	87537695 (2563, 2564) Protein sin. GBans 19328190 (AF074266) - proto-	20466305 (2565, 2566) Nucyerie Art (Nus musculus) 20466305 (2565, 2566) Novel Protein Sim. GBanta gij3261721jembjCAB07057] - (709770) Nucyerie Diotest Brocker (1709770) Nucyerie Diotest Brocker (1709770) Nucyerie Brocker (1709770) Nucyer	(C22/10) hypotherical protein KV0133C [Mycobactenum tuberculosis]	Z0636225 (2567, 2568) Novel Protein sim. GBank gi 3929022 (AF057696) - LspB [Haemophilus ducreyi]	80427330 (2569, 2570) Novel Protein sim. GBank gi 417154 sp P33126 HSB2_ORYSA · HEAT SHOCK PROTEIN 82			95338101 (2575, 2576) Novel Protein sim. GBank gi[5363510]gb AAD42161.1 AF08891 - (AF088916) emiin precursor [Homo sapiens]		1290 19526027 (2579, 2580) Novel Protein sim. GBank gil1169995 sp P46023 GPCR_LYMST - G-PROTEIN COUPLED RECEPTOR GRL101 PRECURSOR	80470266 (2581, 2582) Novel Protein sim. GBank gij2072674 emb CAB08305 - (295120) rhlE [Mycobacterium tuberculosis]	94723316 (2583, 2584) Novel Protein sim. GBank gil1835755 (U86338) - zinc finger Contains protein domain (PF01530) - transcriptfactor Zinc finger, C2HC type		82125908 (2587, 2588) Novel Protein sim. GBank gi[2129173 pir F64453 - oxaloacetate decarboxylase (EC 4.1.1.3) alpha subunit - Methanococcus jannaschii	11686851 (2589, 2590) Novel Protein sim. GBank gi 5441779 emb CAB46803.1
94323988 (2561, 2562)	87537695 (2563, 2564)	20466305 (2565, 2566)		ZUB36325 (2567, 2568)	80427330 (2569, 2570)	20465254 (2571, 2572)	80417530 (2573, 2574)	95338101 (2575, 2576)	11813647 (2577, 2578)	19526027 (2579, 2580)	80470266 (2581, 2582)	94723316 (2583, 2584)	80067536 (2585, 2586)	82125908 (2587, 2588)	11686851 (2589, 2590)
1281	1282	1283					1287	1288	1289				1293		1295

1206	11687004 (2504 2502)				
1207		70630300 (2503 2504) Name Destriction Comment		UNCLASSIFIED	264591, 264639
67	_	Novel Protein sim. GBank	Contains protein domain (PF01367) - polymerase	polymerase	264693
		gif+902 i 9 igop/A-D30000. i PEUU18U - (AEUU18U5) DNA- directed DNA polymerase [Thermotoga maritima]	5-3 exonuclease		
1298		94239506 (2595, 2596) Novel Protein sim. GBank gil1943770 (U97191) - F53F10.1 gene product [Caenorhabditis elegans]		struct	18108348, 265017
1299					264488, 264906, 264909, 22279002, 264566
138		80064867 (2599, 2600) Novel Protein sim. GBank gij3445181 (AC005498) - R31665_2 [Homo sapiens]	Contains protein domain (PF01352) - transcriptfactor KRAB box	transcriptfactor	264605
1301		17939614 (2601, 2602) Novel Protein sim. GBank gil4062973 dbj BAA36204.1 - (AB017138) alpha subunit of malonate decarboxylase [Pseudomonas putida]		UNCLASSIFIED	264906
1302	95416198 (2603, 2604)	_			85658542 285020
1303					2555574, 255540 264908
1304	79377196 (2607, 2608)			UNCLASSIFIED	264508
1305	19905899 (2609, 2610)				264566
1306	13069230 (2611, 2612) Novel Protein sim. G (292669) hypothetica tuberculosis	Novel Protein sim. GBank gij3242273 emb CAB07017 - (292669) hypothetical protein Rv0236c [Mycobacterium tuberculosis]		UNCLASSIFIED	264636
1307				UNCLASSIFIED	264907, 264592, 264764
1308		Novel Protein sim. GBank gi[1502421 (U59433) - 3-ketoacyl Contains protein domain (PF00516) - reductase acyl carrier protein reductase [Bacillus subtilis]	Contains protein domain (PF00516) - Envelope glycoprotein GP120	reductase	264555
1309		79263011 (2617, 2618) Novel Protein sim. GBank gilg5819 pir S16298 - ferric enlerobactin transport protein fepC - Escherichia coli		transport	264906, 18108354
1310		Novel Protein sim. GBank gi[5459220 emb CAB48893.1] - (AL096837) putative iron-sulfur protein [Streptomyces coelicolor A3(2)]		UNCLASSIFIED	264605
1311					35696286, 29331827, 264908, 265008, 264764, 254766, 264686, 21906767, 21906769000000000000000000000000000000000
1312	88061720 (2623, 2624) Novel Protein sim. Gi (AF125158) zinc fingr sapiens]	Novel Protein sim. GBank gil4455118 gb AAD210B4 - (AF125158) zinc finger DNA binding protein 99 [Homo sapiens]	Contains protein domain (PF00096) - dna_rna_bind Zinc finger, C2H2 type	dna_rna_bind	22278995, 22278998, 22278999, 264905, 264908, 265011, 265017, 265019, 264687, 29906768, 265020, 265021, 33657023, 24500, 265021, 23657023, 24654
1313	91225458 (2625, 2626)	91225458 (2625, 2626) Novel Protein sim. GBank gll4929733gbJAD34127.1 AF15189 - (AF151890) CGI-132 Ribosomal protein S16 protein Homo saniens1	Contains protein domain (PF00886) - ribosomalprot Ribosomal protein S16	ribosomalprot	22278996, 22278999, 264259, 20281099, 29146498, 264508, 26512502, 2651659, 26512502, 26513502, 26512502, 26513502, 26512502, 2
<u> </u>					04432350, 04433436, 763011, 263017, 264683, 264288, 21906765, 21906767, 21948627, 21906768, 35695917, 256021, 33657103, 33457109, 18408477, 41408477
					35695855, 60432113, 22279000, 264563, 18108390
1314	_				264693
1315		84357192 (2629, 2630) Novel Protein sim. GBank gi 2589223 (AF026565) - ring finger protein [Mus musculus]	Contains protein domain (PF00097) - Interleukinrecept Zinc finger, C3HC4 type (RING finger)		264691

kinase 56182575, 56181686, 20281171, 29331822, 29331824, 60424269, 29331825, 35696052, 25644045, 264951 60432229, 265018, 265019, 55811150, 56181562, 21906765, 21906768, 35695917, 60170615, 33657023, 65274620, 33657109, 35695763, 3569555, 18108387, 87168518, 60432113, 22279002, 264564	UNCLASSIFIED 264093, 264906, 264909, 264369, 264684	UNCLASSIFIED 18108348, 18108348, 265011, 265017, 18108362, 18108362, 56182323, 18108385, 22279000	264908, 264909, 265006, 265008, 264592, 265019, 264766, 66181562, 18108368, 264629, 18108377, 264636	UNCLASSIFIED 35696286, 55812038, 265018, 21906768, 265020, 263978, 22279002	inscriptfactor 60432049, 29331828, 264907, 264908, 264909, 264910, 55812038, 264601, 264762, 264768, 264768, 264768, 264628, 18108374, 264634, 264635, 18108385	ia_ma_bind 29331828, 264908, 265020, 33657023, 264693, 264404	synthase 65274572, 56994075, 264259, 29331822, 29331827, 264104, 56182435, 87168474, 18108351, 264288, 21906766, 21906767, 35695917, 265020, 264693, 65274791, 56182323, 18108387	22278996, 22278998, 29331828, 264905, 264907, 29331830, 264908, 264510, 265008, 264595, 264769, 21906754, 265018, 264288, 264768, 264631, 264631, 264632, 264636, 264564, 264565, 264566	35696286, 264905, 264906, 264907, 264908, 264909, 264909, 264910, 264593, 33657402, 264758, 85658542, 264760, 264768, 264769, 264691, 35696423	UNCLASSIFIED 264259, 66712502, 264682, 264683, 264635
<u> </u>	'n	in		5	Contains protein domain (PF00651) - transcriptfactor BTB/POZ domain	Contains protein domain (PF00096) - dna_ma_bind Zinc finger, C2H2 type		Contains protein domain (PF01852) - START domain	Contains protein domain (PF00090) - protease Thrombospondin type 1 domain	<u> </u>
1316 95361609 (2631, 2632) Novel Protein sim. GBank gi 5689407 db BAA82997.1 - (AB028958) KIAA1035 protein [Homo sapiens]	Novel Protein sim. GBank giļ4836757]gbJAAD30541.1[AF13491 - (AF134918) semaphorin subclass 4 member G [Mus musculus]	95322893 (2635, 2636) Novel Protein sim. GBank gil4680204[gb AAD27567.1 AF11417 - (AF114171) hypothetical protein [Sorghum bicolor]		86603567 (2639, 2640) Novel Protein sim. GBank gile240183 dbj BAA74870.1	86676351 (2641, 2642) Novei Protein sim. GBank giļ4886505 emb CAB43377.1{	87755272 (2643, 2644) Novel Protein sim. GBank gi[5262591 emb CAB45736.1	94845931 (2645, 2646) Novel Protein sim. GBank gi 5459516 db BAA82407.1 - (AB029821) phosphatidylethanolamine N-methyltransferase [Homo sapiens]	Novel Protein sim. GBank gij5031717Iref NP_005704.1 pGPBP - goodpasture antigen- START domain binding protein	1325 94847471 (2649, 2650) Novel Protein sim. GBank gij3294501 (U64857) - similar to the DPTUKunitz family of inhibitors; most similar to tissue factor pathway inhibitor precursor [Caenorhabditis elegans]	87316289 (2651, 2652) Novel Protein sim. GBank gi 1397275 (U61947) - C06G3.8 gene product [Caenorhabdilis elegans]
95361609 (2631, 263 <u>2)</u>	88055167 (2633, 2634)	95322893 (2635, 2636)	94238546 (2637, 2638)		86676351 (2641, 2642)			87737614 (2647, 2648)	94847471 (2649, 2650)	
1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326

1327	95322897 (2653, 2654)	95322897 (2653, 2654) Novel Protein sim. Gbank gil728832Ispl⊅39189IALUZ_HUMAN - IIII ALU SUBFAMILY SB VVARNING ENTRY IIII	Contains protein domain (PF00279) - UNCLASSIFIED Plant lipid transfer protein family		18108398, 22278996, 22278997, 22278999, 264091, 264259, 29331822, 29331824, 29331825, 29331827, 29331828, 264005, 264905, 564905, 564905, 564005, 265009, 21906754, 265010, 265011, 265017, 265019, 264681, 26448, 264764, 264684, 264288, 264685, 265000, 264691, 264691, 264691, 264692, 265000, 264690, 264691, 264691, 264691, 264692, 264690, 264690, 3957349, 18108370,
1328	87753493 (2655, 2656)			UNCLASSIFIED	18108374, 55810764, 35695855, 264634, 56182323, 83373044, 87168518, 60432113, 22279000, 22279002, 264563, 264567, 264488, 22278997, 29331826, 264585, 264682, 264766, 22279002, 264482, 264682
1329		87755276 (2657, 2658) Novel Protein sim. GBank gil4678224[gb AAD26669.1 AC00713 - (AC007135) unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	204-097 22278996, 29331827, 264684, 264692, 33657109
1330		87727737 (2659, 2660) Novel Protein sim. GBank gi 437310 (L23504) - nodulin [Medicago truncatula]		UNCLASSIFIED	264259, 29331825, 264512, 265019, 265021, 264555, 264558, 56526486
1331		87376764 (2661, 2662) Novel Protein sim. GBank gil4589586 dbj BAA76815.1 - (AB023188) KIAA0971 protein [Homo sapiens]		UNCLASSIFIED	264259, 29331826, 29331827, 35696052, 29331828, 60170831, 264448, 264686, 21906765, 55811957, 265020, 33657023, 33657109, 263973, 55811576, 35696423, 35695855, 56182323
1332	94845937 (2663, 2664)	94845937 (2663, 2664) Novel Protein sim. GBank gi[3459516 dbj BAA82407.1 - (AB029821) phosphatidylethanolamine N-methyltransferase [Homo sapiens]		synthase	65274572, 22278996, 56994075, 22278999, 60432049, 264259, 29331822, 29331826, 60432289, 29331827, 35696052, 52644045, 56182435, 224510, 21906754, 87168559, 264686, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 23857023, 56182323, 18108387, 60432113, 2279002
1333		88098476 (2665, 2666) Novel Protein sim. GBank gi 5689527 db BAA83047.1 - (AB029018) KIAA1095 protein [Homo sapiens]		UNCLASSIFIED	60432289, 66712502, 264591, 60433356, 60433438, 55812038, 265010, 264639, 56526486
1334	87592388 (2667, 2668)	87592388 (2667, 2668) Novel Protein sim. GBank gil 2662536 (AF036685) - Similar to protein-tyrosine phosphatase [Caenorhabditis elegans]		phosphatase	264905
1335	87644798 (2669, 2670)	1335 87644798 (2669, 2670) Novel Protein sim. GBank gil4240285 dbj BAA74921.1 - (AB020705) KIAA0898 protein [Homo sapiens]	Contains protein domain (PF00643) - UNCLASSIFIED B-box zinc finger.	UNCLASSIFIED	22278998, 22278999, 29331827, 264509, 264511, 265007, 265008, 265009, 60433438, 21906754, 87168559, 265017, 264288, 21906765, 21906768, 21906769, 265020, 33657109, 27486264, 18108374, 264556, 264638, 264557, 60170394, 264559, 18108385, 264563

264509, 264905, 264512, 264764, 264693, 264635, 264637	56182575, 56994075, 22278998, 22278999, 264092, 264259, 60432289, 29331826, 264906, 264908, 264908, 264112, 265008, 265009, 60433356, 55812038, 33657084, 265011, 265017, 265018, 265019, 264682, 264448, 264683, 264369, 21906766, 21906769, 18108379, 55811576, 27486261, 20281069, 18108379, 55811576,	29331822, 265010, 264288, 264689, 1810827, 263082, 263082, 26308, 2646890, 264689, 264689, 264689, 264689, 264689, 264689, 264689, 264	35696052, 264909, 264688, 264556, 264558	264905, 264907, 87168559, 264764	264681, 264685, 264686, 264692	264629	264910, 264686, 264534	263978	264909, 60170394	22278998, 264907, 264681, 264685, 264689, 265020, 264693, 22279000, 22279002, 264566	264906, 264908, 264909, 265006, 264910, 265011, 265017, 264764, 264766, 264767, 264769, 264631, 264634, 264638, 264567, 264486	264488, 264489, 264509, 264510, 264511, 264512, 264591, 264592, 264601, 264684, 264685, 264769, 264532, 264534, 264555, 264556, 264557, 264558, 22279002, 264586
	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED			tm7	struct		gtycoprotein	kinase	glycoprotein	UNCLASSIFIED
							Contains protein domain (PF00560) - struct Leucine Rich Repeat			Contains protein domain (PF00069) - kinase Eukaryolic protein kinase domain	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat	
1336 87787890 (2671, 2672) Novel Protein sim. GBank gij465445 sp P33485 VNUA_PRVKA - PROBABLE NUCLEAR ANTIGEN	94312042 (2673, 2674) Novel Protein sim. GBank gi[5689471 dbj BAA83019.1 - (AB028990) KIAA1067 protein [Homo sapiens]		Novel Protein sim. GBank gi 1176422 (U43194) - rhophilin Mus musculus			80089017 (2683, 2684) Novel Protein sim. GBank gi 5019564 emb CAB44507.1 - (AL035542) dJ994E9.5 (hs6M1-17 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) protein)) [Homo sapiens]	Novel Protein sim. GBank gil4557543 ref NP_001384.1 pECM2 - extracellular matrix protein 2		91225546 (2689, 2690) Novel Protein sim. GBank gij2144101 pir 55210 - tricarboxylate carrier - rat (fragment)	80255717 (2691, 2692) Novel Protein sim. GBank gil3881052 emb CAA19523 - (AL023843) predicted using Genefinder; similar to serine/threonine kinase; cDNA EST yk246a12.3 comes from this gene; cDNA EST yk358c10.5 comes from this gene; cDNA EST EMBL:M89047 comes from this gene; cDNA EST EMBL:M89047 comes from this gene; cDNA	1347 80417393 (2693, 2694) Novel Protein sim. GBank gil4504379 ref NP_003658.1 pHG38 - orphan G protein- coupled receptor HG38	87352335 (2695, 2696) Novel Protein sim. GBank gij3399720 dbj BAA32100 - (AB010999) peptidylarginine deiminase type IV [Rattus norvegicus]
87787890 (2671, 2672)	94312042 (2673, 2674)	80366114 (2675, 2676)	80249231 (2677, 2678) Novel Protein sim. [Mus musculus]	88316311 (2679, 2680)	86101485 (2681, 2682)	80089017 (2683, 2684)	80082862 (2685, 2686) Novel Protein sim. gil4557543 ref NP_ protein 2	20562559 (2687, 2688)	91225546 (2689, 2690)	80255717 (2691, 2692)	80417393 (2693, 2694)	87352335 (2695, 2696)
1336	1337	1338	1339	7			1343		1345	1346	1347	1348

2278999, 264259, 22278996, 22278998, 22278999, 264259, 29331822, 29331824, 35696052, 264509, 264905, 264906, 264906, 264906, 264906, 264909, 264909, 264511, 265006, 264512, 264910, 60170831, 264591, 60433438, 264767, 21906754, 265017, 255018, 264605, 21906765, 21906767, 21906768, 21906768, 21906768, 21906768, 21906768, 21906768, 21906768, 21906768, 21906768, 21906768, 21906768, 264691, 264692, 33657023, 264693, 33657349, 18108374, 18108376, 35596423, 264564		ED 264906, 264907, 264638	ED 264693, 263981		264512, 264910, 264758, 264596, 55811386,	265011, 264605, 55811150, 264762, 264764, [264766, 52644229, 56181562, 35695917]	265022, 33657023, 264693, 35695763,	60431528, 264629, 263978, 35696423,	35695855, 264630, 264634, 264635, 264636. 264637, 264638, 264639, 18108385, 264563,		ED 22278995, 22278999, 29331826, 264906, 266008, 2366740, 24006764, 266044	87168559, 264684, 264369, 264769, 264689,	21906765, 21906768, 52644150, 33657023,	264692, 264693, 18108374, 83373044,	22278996, 22278997, 264259, 66714117.	264511, 21906754, 265010, 264769, 264689,	21906765, 21906768, 21906769, 264532,	27486262, 264629, 264636, 264556, 264638, 264639, 264482, 26484
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED								UNCLASSIFIED				dna rna bind	1		
						-									Contains protein domain (PF00076) - dna ma bind	RNA recognition motif. (a.k.a. RRM,	RBD, or RNP domain)	
Novel Protein sim. GBank gi[2144101 pir 55210 - tricarboxylate carrier - rat (fragment)		Novel Protein sim. GBank gil4887239 gb AAD32246.1 - (AF064564) BAW protein [Fugu rubripes]		Novel Protein sim. GBank gi[2144101 pir 155210 - tricarboxvlate carrier - rat (fragment)						0.6	Novel Protein Sim. GBank qi 4689108 ab AAD27763 1 AF07703 - (AF077030)	hypothetical 43.2 kDa protein [Homo sapiens]			Novel Protein sim. GBank gil1469199 dbj BAA09487 -	The KIAA	sapiens]	
1349 91225548 (2697, 2698), Novel Protein sim. (tricarboxylate carrie	1350 87093136 (2699, 2700)	87361327 (2701, 2702) Novel Protein sim. (AF064564) BAVV	80076386 (2703, 2704)	95345417 (2705, 2706) Novel Protein sim. tricarboxylate carrie						10000	93330843 (2707, 2708) Novel Protein sim. 93330843 (2707, 2708) 93350843 (2707, 2708)				88260186 (2709, 2710) Novel Protein sim. (
0.46 0.46	1350	1351	1352	1353						1	2 2 3 4				1355			

1356	1356 95313991 (2711, 2712) Novel Protein sim	Novel Protein sim GBank oil1113865 (1140342), pinein		to the contract of	18108207 2228006 22228008 2228008
		[Mus musculus]			264094, 29331829, 252, 1522, 16339, 264094, 29331829, 254905, 265006, 265007, 265008, 265010, 265017, 265018, 265019, 264764, 18108354, 264689, 21906765, 265022, 18108364, 35696423, 83373044, 18108387
1357		88260268 (2713, 2714) Novel Protein sim. CBank gil897893 emb CAA90330 - (Z50026) phosphatidylcholine transfer protein [Bos taurus]	Contains protein domain (PF01852) - START domain		264259, 29331822, 29331825, 264510, 8716859, 265018, 264448, 264288, 21906765, 21906766, 21906768, 265021, 244693, 18108376
1358		Novel Protein sim. GBank gi[556219 (L36831) - transcription regulator [Mus musculus]			264757
1359				UNCLASSIFIED	264907, 264909, 264510, 264511, 264512, 18108351, 26454, 33657023, 18108374, 264634, 264635, 264638, 264639, 264567
1360	87738272 (2719, 2720)	87738272 (2719, 2720) Novel Protein sim. GBank gi[2598282 emb CAA75612 - (Y15417) acetateCoA ligase (Coprinus cinereus)		synthase	60432289, 264605
1361	87593527 (2721, 2722)	87593527 (2721, 2722) Novel Protein sim. GBank gil5689443 dbj BAA83005.1 -	Contains protein domain (PF00536) - UNCLASSIFIED	UNCLASSIFIED	35696286, 22278997, 22278999, 264259,
		(AB028976) KIAA1053 protein [Homo sapiens]	SAM domain (Sterile alpha motif)		29331826, 264508, 264509, 264905, 264907, 264908, 265007, 265009, 33109954,
					21906754, 87168474, 265011, 264761,
					264683, 264288, 264766, 264769, 264689.
					z 1900/06, zejozu, zejuz 1, jsej 70zs, 55811576, 35696423, 264634, 60432113.
					22279002, 264482, 264486
1362	95287961 (2723, 2724)	32989.11 -	Contains protein domain (PF00400) - eph	hde	56182575, 56181686, 60432049, 264259,
		(AB028960) KIAA1037 protein [Homo sapiens]	WD domain, G-beta repeat		29331822, 56182181, 29331827, 35696052,
					29331828, 264905, 264906, 264908, 264595, 55812038, 85658542, 55811150, 264681
					264288, 264369, 56181562, 60431528,
,	2010, 0110,0110				55810764, 35696423, 60431850, 264558
င်ရင် ရင်	63/384/6 (2/25, 2/26)	GBank gij1130494 (U35776) - ADP- 1-directed GTPase activating protein	Contains protein domain (PF01412) - UNCLASSIFIED Putative GTP-ase activating protein	UNCLASSIFIED	264488, 29331826, 264907, 264687, 264689, 264693
į	\neg	Inditions notivegrous,	tor Art		
1364	88179488 (2727, 2728)				60432289, 60433356, 60433438, 87168559, 264603, 18108351, 21906766, 35696423, 60432113
1365	83003108 (2729, 2730)	83003108 (2729, 2730) Novel Protein sim. GBank gil4589562 dbj BAA76803.1 - (AB023176) KIAA0959 protein [Homo saplens]		oncogene	264766
1366	87003262 (2731, 2732)		Contains protein domain (PF00153) - transport		265007
		hypothetical protein YPR021c - yeast (Saccharomyces cerevisiae)	Mitochondrial carrier proteins		

1369		1367 87721210 (2733, 2734) Novel Protein sim. GBank gil464561 spil-35289 RB15_RAT (AL050019) hypothetical protein [Homo sapiens] (AL050019) hypothetical protein [Homo sapiens] (AL05019) Novel Protein sim. GBank gil464561 spil-35289 RB15_RAT (AL050019) Novel Protein sim. GBank gil2062702 (U90550) - (AL050018) Novel Protein sim. GBank gil2062702 (U90550) - (AL050018) Dulyrophilin [Homo sapiens] (AL050019) Novel Protein sim. GBank gil2062702 (U90550) - (AL050019) Novel Protein sim. GBank gil2062702 (U90550) -	Contains protein domain (PF01342). Contains protein domain (PF00071). Ras family	SIFIED IFFED Channel	264488, 52646842, 52646365, 22278995, 56994075, 35696286, 22278996, 22278999, 264259, 29331824, 66714117, 29331825, 60432289, 35696052, 264905, 264905, 264905, 264905, 264907, 264908, 264909, 265009, 60170831, 33657402, 256110, 265009, 60170831, 33657402, 55612038, 21906754, 265011, 26718559, 264910, 265009, 60170831, 264689, 264682, 264683, 264288, 264368, 264689, 264689, 264689, 264689, 264689, 264689, 264689, 264689, 264689, 264689, 264689, 264689, 264689, 264689, 264689, 264690, 264691, 18108362, 33657023, 33657109, 33657349, 264628, 18108370, 18108374, 18108374, 18108374, 18108374, 18108374, 18108374, 18108374, 18108375, 264555, 264637, 265007, 265029, 264690, 264691, 264690, 264691, 264690, 264691, 2646
1371	95336512 (2741, 2742)	95336512 (2741, 2742) Novel Protein sim. GBank gi 5032203 ref NP_005714.1 pTSPA - tetraspan 5	Contains protein domain (PF00335) - (4 transmembrane'segments integral membrane proteins	glycoprotein	264404, 264563, 264566, 264486 22278996, 264269, 29331622, 29331824, 29331825, 29331826, 29331827, 264905, 264509, 29331830, 66712502, 265008, 265009, 264758, 33657084, 86558542, 265010, 265018, 265019, 264762, 26448, 35695817, 33657109, 33657182, 33657349, 36595855, 264558, 22279002, 264663

1372	80248517 (2743, 2744)	1372 80248517 (2743, 2744) Novel Protein sim. GBank gill840708idbill8AA093341 -		collagen	263978
		(D50685) trans-stalidase [Trypanosoma cruzi]			
1373	80499421 (2745, 2746)			UNCLASSIFIED	264769, 21906765, 21906767, 22278999, 264691, 264910, 55812038, 265010, 264681, 264684
1374		Novel Protein sim. GBank giļ111876 pir JC1241 - beta- interferon-induced protein - rat		interferon	264907, 264510, 265007, 264512, 265008, 6043229, 264689, 65274791, 264555, 264567, 83373044, 60432113
1375	94236942 (2749, 2750) Novel Protein sim. (AF051155) G beta	Novel Protein sim. GBank gil5649176 gblAAD03500.2 - (AF051155) G beta-like protein GBL [Rattus norvegicus]	Contains protein domain (PF00400) - kinase WD domain, G-beta repeat	kinase	52644507, 52645156, 52646842, 52646365, 56182575, 56181686, 22278996, 56994075, 35696286, 22278996, 56994075, 35696286, 22278997, 22278998, 22278999, 22278999, 22278999, 22378997, 22278998, 22278999, 29331824, 29331825, 29331824, 29331826, 29331824, 29331826, 29331826, 29331827, 263019, 264906, 264906, 264906, 264906, 264906, 264906, 264906, 264909, 265009, 33657402, 60433356, 52646317, 33109954, 33957402, 264448, 264764, 264288, 264369, 264488, 264764, 26448, 264764, 265020, 265021, 52644150, 33657023, 264693, 264634, 264634, 264634, 264634, 264636, 264636, 264636, 264656, 264636, 264656, 264666, 2
1376	87399050 (2751, 2752)			UNCLASSIFIED	264768, 264769, 35695917, 22278997, 264691, 264259, 29331822, 264693, 35696052, 264508, 264509, 264693, 264606, 264608, 264608, 264610, 264611, 264611, 264631, 265008, 264634, 264634, 264632, 264639, 264519, 264509, 264518, 264637, 264637, 264637, 264637, 264639, 264588, 2652010, 265011, 264600, 264563, 264564, 264565, 264564, 264565, 264656, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 264665, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465, 26465,
1377	86964242 (2753, 2754) Novel Protein sim. granule ATPase II	Novel Protein sim. GBank gi[1663648 (U75321) - chromaffin Contains protein domain (PF00122) - ATPase_associated granule ATPase II homolog [Mus musculus]	n Contains protein domain (PF00122) - E1-E2 ATPase	ATPase_associated	

22278994, 22278998, 264093, 264094, 264259, 29331824, 29331827, 265009, 265018, 265019, 18108351, 264369, 264288, 29148627, 55811957, 264691, 18108366, 33657109, 18108368, 264635, 263981, 18108385	265020	264510, 264512, 265009, 264288, 264564	87168559, 265017, 264628, 22279002	22278994, 22278997, 22278999, 29331822, 29331824, 29331826, 265007, 60432229,	60433356, 85656542, 265017, 265018, 264685, 264768, 21906766, 35695917, 33657023, 27486261, 27486262, 35695763, 35695855, 87168518, 22279002	18108396, 264692	264488, 264508, 264509, 264905, 264906,	264908, 264909, 264511, 264512, 264910, 264760, 18108351, 264766, 264769,	35695855, 264630, 264636, 264555, 264638, 264483, 264564, 264486	35696052, 55811386, 264688, 21906765,	20022) 0000 020 020 064007	56182435, 28531026, 264800, 204807.	33657023, 83373044, 264566	264591	65274572, 22278999, 264259, 29331826,	[29331827, 35696052, 264509, 264907, 264908, 264908, 264909, 265006, 265008, 60170831,	33657402, 60433438, 264596, 21906754,	87168559, 264600, 265017, 264683,	18108354, 52644229, 21906765, 21906766,	21906767, 21906768, 21906769, 265021,	264692, 33657023, 33657109, 35695855, 264558, 60170394, 83373044, 22279000
		nucl_recpt	UNCLASSIFIED	UNCLASSIFIED		UNCLASSIFIED	UNCLASSIFIED			UNCLASSIFIED	,50	<u> </u>		UNCLASSIFIED	glycoprotein						
				Contains protein domain (PF00583) - UNCLASSIFIED Acetyltransferase (GNAT) family																	
1378 87595071 (2755, 2756) Novel Protein sim. GBank gil+107015(db) BAA36293 - (AB001772) PEM-5 [Ciona savignyi]	85679344 (2757, 2758) Novel Protein sim. GBank gi 3252872 (AF035620) - BRCA1- associated protein 2 (Homo sapiens)	87627962 (2759, 2760) Novel Protein sim. GBank gil4837737]gb AAD30662.11- (AF096834) germ cell specific Y-box binding protein [Homo saplens]	88179656 (2761, 2762) Novel Protein sim. GBank gil4731580jgb[AAD28508.1 AF12538 - (AF125384) L82A [Drosophila melanogaster]							86378788 (2769, 2770) Novel Protein sim. GBank gi[2384732 (AF015911) - NAC-1	protein [haitus ilorvegicus]	91013049 (2771, 2772) Novel Protein sim. GBank gil2384910 (AF-022962) - contains similarity to the A-type potassium current class of	channel proteins [Caenorhabditis elegans]	87797958 (2773, 2774) Novel Protein sim. GBank gil4160304[emb CAA10600] - [(AJ132192) HS1 binding protein 3 [Mus musculus]	95101652 (2775, 2776) Novel Protein sim. GBank	gil4895164 gb AAD32753.1 AC00723 - (AC007231) putative disease resistance protein [Arabidoosis thaliana]					
87595071 (2755, 2756)	_	87627962 (2759, 2760)	88179656 (2761, 2762)	94847576 (2763, 2764)		87860598 (2765, 2766)	1384 86915895 (2767, 2768)			_	_	91013049 (2771, 2772)		87797958 (2773, 2774)	95101652 (2775, 2776)						
1378	1379	1380	1381	1382		1383	1384			1385	3	1386		1387	1388						

31822, 3331828, 12038, 1, 264369, 1906768, 3, 5811576, 4,	45080, 44045, 68474, 0, 264682, 2, 5695917, 150, 5274791, 518,	64448. 555,	2278998, 45080, 9331826, 3656970, 77, 265008, 1677, 1767, 1767, 109, 109, 109, 109, 109, 109, 109, 109		
65274572, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331828, 29146499, 264906, 66712502, 55812038, 265017, 265018, 18108351, 264369, 265020, 265021, 18108371, 264056, 21906766, 21906766, 21906766, 21906766, 21906766, 21906768, 264552, 264557, 60170394, 264555, 264556, 264557, 60170394, 222730000, 264563, 264564	52645156, 52646365, 264259, 52645080, 29331825, 29331826, 264806, 52644045, 265009, 33657084, 52644296, 87168474, 87168559, 265017, 265018, 264760, 264682, 264429, 21905765, 21906769, 365020, 265021, 60170615, 52644150, 33657023, 27486261, 27486264, 65274791, 264631, 264555, 52644332, 87168518, 22279000, 264567	29331826, 29331828, 29331830, 264448, 264288, 33657023, 18108365, 264555, 264556, 83373044	5564642, 65274572, 22278994, 22278995, 325696246, 66994075, 22278994, 22278999, 60432049, 264259, 52645080, 22278999, 60432049, 264259, 52645080, 29331822, 29331824, 60432289, 29331826, 29331827, 35696052, 29331828, 33656970, 264907, 52644045, 265006, 265007, 265007, 2644296, 265010, 87169559, 265017, 264469, 265019, 264468, 264369, 265019, 264468, 264369, 265019, 264468, 264369, 265019, 264468, 264626, 27486262, 27484, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27484, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 27486262, 274862626, 27486262, 2748	264763, 264631	265009, 18108381
UNCLASSIFIED	peptidase			UNCLASSIFIED	UNCLASSIFIED UNCLASSIFIED
Contains protein domain (PF00641) - UNCLASSIFIED Zn-finger in Ran binding protein and others.		Contains protein domain (PF00709) - Adenylosuccinate synthetase		Contains protein domain (PF00560) - UNCLASSIFIED Leucine Rich Repeat	
1389 91256016 (2777, 2778) Novel Protein sim. GBank gijs689387 dbj BAA82977.1 - (AB028948) KIAA1025 protein [Homo sapiens]	94111916 (2779, 2780) Novel Protein sim. GBank gij3702295 (AC005783) - R33083_1 [Homo sapiens]	91227345 (2781, 2782) Novel Protein sim. GBank gi 1346910 sp P28650 PUA1_MOUSE - ADENYLOSUCCINATE SYNTHETASE, MUSCLE ISOZYME (IMPASPARTATE LIGASE)	94311097 (2783, 2784) Novel Protein sim. GBank gil726286 (U22394) - mSin3A [Mus. musculus]	(95)	15028819 (2787, 2788) 95361471 (2789, 2790) Novel Protein sim. GBank gi[2274845 dbj BAA21534 - [D88461) N-WASP [Rattus rattus]
91256016 (2777, 2778)		91227345 (2781, 2782		80409472 (2785, 2786)	
1389	1390	1391	1392	1393	1395 1395

1396	1396 95363253 (2791, 2792) Novel Protein sim. Gl	Novel Protein sim. GBank gi[2135904[pir] 54810 - pH			בהניבההם הזהנים ההחמדההם דההמדפהם	Γ
		E1F1 - human			444, 444, 644, 644, 644, 644, 644, 644,	
_					60432289, 29331828, 29146498, 29146499,	
					264907, 264908, 29331830, 264909, 265006,	
					265007, 265008, 265009, 60433356, 265010,	
					264602, 265017, 265018, 265019, 18108354,	
					52644229, 18108358, 21906767, 29148627.	
					21906768, 21906769, 29148629, 29148784,	
	-				265021, 265022, 18108368, 18108374,	_
1307	87624247 /0702 0704)				56182323, 18108385, 264563, 264567	
2				UNCLASSIFIED	35696286, 264907, 66712502, 264510,	Т
1308	_				35695917, 264692, 264693, 35696423	_
2	1233007 (2733, 2730) Novel Protein Sim. Gt	Novel Protein Sim. GBank gil5420389 emb CAB46680.1 -			264259, 29331822, 29331824, 29331825,	Т
		(A3243460) proteopnosphoglycan [Leishmania major]			29331827, 35696052, 33656970, 87168474,	
					265018, 265019, 264682, 264768, 21906767.	_
			-		265020, 33657023, 27486261, 55811576,	
					264632, 264639, 83373044, 87168518,	_
1300	_	87631076 (2707 2700) Marris Dadain Ch. 1			22279002	
<u>}</u>		gi[2496887]spj@09232 YQ22_CAEEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III		UNCLASSIFIED	264768, 18108370, 264555, 264557	T
1400		95419064 (2799, 2800) Novel Protein sim. GBank gi 283920 pir S27939 - tensin -	Contains protein domain (PF00017) - UNCLASSIFIED	UNCLASSIFIED	56182575 22278994 22278997 264259	1
		chicken	Src homology domain 2		2012E013, EEE10034, EEE10337, 204233,	
					2331828 264908 46182434 264112	-
					265000 265011 265017 265010 265010	_
					200009, 200011, 200017, 200018, 200019,	-
					264750, 264762, 264765, 264288, 264685,	_
					264687, 56181562, 264769, 21906766,	_
					21906767, 55811957, 264691, 264692,	_
	_				264628, 264629, 55811576, 264634, 264555,	
_					264637, 264557, 264638, 18108381, 264558,	
1401	91226179 (2801 2802)	Partial Distriction of the Control o			18108384, 60432113, 22279000	
2	21220313 (2001, 2002)	3 (2001, 2004) Novel Protein Sim. GBank gij3256185jemb CAA15485 -	Contains protein domain (PF00790) - UNCLASSIFIED		65274572, 60432289, 264909, 264758,	Т
1400	_	(ALCUdes) application (Home saplens)	VHS domain		264768, 21906769, 22279002	_
<u> </u>		33301473 (2003, 2004) [Novel Pfotein sim. GBank gi 1515427 (U57523) - nel	Contains protein domain (PF00008) - tgf		264905, 264907, 264908, 264909, 264112,	Т
1402		liomolog (Homo sapiens)	EGF-like domain		264693, 33657109, 264634	
3		34 14 1933 (2003, 2000) Novel Protein sim. GBank gi[5262615]emb[CAB45747 1]			65274572, 66712502, 265017, 264448,	Т
		(ALugu Ibo) nypotnetical protein [Homo sapiens]			264288, 21906765, 21906769, 264693,	
1404	90935393 (2R07 2R08)				55811576, 65274791, 60432113	
<u> </u>	(2002, 2007)			UNCLASSIFIED	65274572, 22278998, 29331822, 29331828,	_
_					66712502, 265008, 60433438, 265017,	
					264693, 18108385	_

26448, 56994075, 35696286, 29331822, 29331826, 29331826, 29331828, 35696052, 264508, 264906, 264907, 264908, 264510, 264511, 264910, 33657402, 264594, 264758, 264600, 264604, 264652, 18108351, 264764, 33657023, 33657109, 264628, 264634, 83373044, 22279002, 264563, 264482, 264486, 264567	collagen 264907, 264605	UNCLASSIFIED 35696286, 22278999, 264094, 264259, 66714117, 29331826, 29331827, 29331828, 29146498, 264107, 264908, 265008, 265008, 264910, 60433438, 265011, 265017, 18108351, 264448, 264288, 264686, 21906765, 21906769, 264628, 264686, 21906765, 21906769, 264628, 263972, 18108374, 35696423, 55811576, 264631, 264557, 264558, 83373044, 18108385, 87168518, 60433113, 22279002	Contains protein domain (PF00008) - oncogene 264488, 264488, 35696286, 264109, 264508, EGF-like domain 264908, 264509, 265008, 265009, 264900, 264908, 265009, 265009, 265009, 264900, 264908, 264908, 265009, 265009, 264901, 33657402, 264757, 264757, 264757, 264758, 265011, 26401, 265017, 264760, 264762, 264681, 264685, 264685, 264692, 264693, 264691, 33657109, 264629, 264629, 35696421, 33657109, 264638, 264639, 264639, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264668, 264568, 264568, 264566,	264693	Contains protein domain (PF00386) - complement 29331826, 264112, 264512, 265009, 265010, 2014 domain 264601, 264686, 264769, 21906767, 263974, 264631, 264566
			Contains protein do		Contains protein do
1405 95095068 (2809, 2810) Novel Protein sim. GBank gil854065[emb CAA58337] - (X83413) U88 [Human herpesvirus 6]	87612369 (2811, 2812) Novel Protein sim. GBank gij624076 gb AAC96425.1 - (U42580) contains Pro-rich Px motifs: SPKPP (20X), PEPPA (9X); similar to soybean pro-rich cell wall protein, corresponds to Swiss-Prot Accession Number P13993 [Paramecium bursaria Chlorella virus 1]	94129872 (2813, 2814) Novel Protein sim. GBank gilz827886 (AF015037) - endooligopeptidase A related protein; EOPA related protein [Oryctolagus cuniculus]	Novel Protein sim. GBank gi 2564953 (AF030001) - unknown [Mus musculus]	66644385 (2817, 2818) Novel Protein sim. GBank gil2662165 db BAA23714 - (AB007902) HH0712 cDNA clone for KIAA0442 has a 574- bp insertion at position 1474 of the sequence of KIAA0442. [Homo sapiens]	86612587 (2819, 2820) Novel Protein sim. GBank gi[2493790]spl060994ACR3_MOUSE - 30 KD ADIPOCYTE COMPLEMENT-RELATED PROTEIN PRECURSOR (ACRP30) (ADIPOCYTE SPECIFIC
95095068 (2809, 2810)	87612369 (2811, 2812)	94129872 (2813, 2814)	95361477 (2815, 2816) Novel Protein sim. Gunknown (Mus musc	66644385 (2817, 2818)	86612587 (2819, 2820)
1405					1410

2821 28221	1411 B7818641 (2821 2822) Movel Bratain aim Coast			
	gil3123155[splP91343]YM3M_CAEEL - HYPOTHETICAL	Contains protein domain (PF00400) - kinasereceptor WD domain, G-beta repeat	kinasereceptor	22278995, 22278997, 22278999, 29331822, 60432280, 20331838, 284067, 266047
	49.0 KD TRP-ASP REPEATS CONTAINING PROTEIN			265019, 264682, 21906767, 21906768,
	TOSTO, SIN CHROMOSOME			21906769, 265020, 264690, 264691,
_				33657023, 33657109, 27486264, 264628,
_				263972, 264634, 264558, 18108385
	04446660 (2025 2026) No. 10 October 10 Contraction		UNCLASSIFIED	264757
	Novel Platein Sim. Gbank gij38/9121jembjCAA94370j -	Contains protein domain (PF00023) - homeobox	homeobox	56994075, 29331822, 35696052, 29331828,
	lankvrin (PIR Acc. No. S32271): «DNA EST EMBL: JOANSE	Ank repeat		29331830, 264909, 52644045, 264510,
	compartion this construction of First construction of the construc			52644296, 85658542, 87168474, 265017,
	from this constitution of the contract of the			265018, 264681, 264687, 21906768,
	THORIT WIS BELIEF, CONA EST EMBL: D32/23 comes from this			35695917, 265020, 52644150, 264692,
	gene, conversor E	•		263967, 27486264, 35695763, 264639,
15				18108387, 264566
2	otologic (2021, 2020) Movel Protein Sim. GBank gil3252981 (AF068921) - Ras- binding protein SUB-8 (Mus musculus)	Contains protein domain (PF00560) - UNCLASSIFIED	UNCLASSIFIED	264682, 264683, 265022, 264636
12	94326948 (2820 2820) Maior Bertain CB-11 (Mars 1922)	Leucine Rich Repeat		
_ `	nover Florent sim. GBank gij1671187 (U90439) - unknown			52646365, 56182575, 22278994, 22278995,
	protein (Arabidopsis thaliana)			56994075, 22278996, 22278997, 22278998.
				22278999, 264259, 29331822, 29331824,
-				29331825, 29331826, 29331827, 29331828,
_				29146498, 66712502, 29331830, 52644045.
_				264113, 264511, 33657402, 264757,
_				21906754, 55811386, 265017, 265018,
_				265019, 264761, 264683, 264369, 264288,
				264686, 264689, 21906766, 21906767,
				29148627, 21906769, 55811957, 265020,
_				265021, 264690, 33657023, 65274620,
_				52645129, 27486262, 27486264, 60431528,
				264629, 35695855, 56182323, 264559,
1				BO420419 984404 99920000 384409

18108392, 18108394, 18108397, 18108398, 22278996, 35696286, 22278997, 22278998, 22278999, 29331822, 35696052, 29331828, 29146498, 264905, 264907, 264908, 264905, 264907, 265007, 265008, 265009, 60170831, 264595, 18108348, 21906754, 87186474, 265019, 265011, 8718659, 265017, 265018, 265019, 264621, 8718659, 265017, 265019, 264621, 18108351, 2644229, 264688, 264682, 264682, 264681, 265020, 21906766, 21906766, 21906766, 21906766, 21906768, 264629, 18108368, 264629, 18108379, 55811576, 65274791, 264634, 264636, 56182323, 18108381, 264634, 264636, 56182323, 18108381,	264107, 264448	264259, 264908, 265010, 52644229, 21906764, 21906768, 264690, 264639, 18108388	264259, 60432289, 265006, 87168474, 264288		18108356, 18108396, 18108397, 21906766, 18108399, 21906767, 56182575, 21906768, 21906769, 56181686, 55811957, 35695917, 35696286, 22278996, 22278997, 22278999, 262021, 265022, 60170615, 264259, 33657023, 29331822, 56182181, 29331822, 56182181, 29331822, 33657109, 29331822, 32857109, 29331829, 264905, 264509, 20281149, 18108370, 264905, 264509, 20281149, 265017, 265007, 60431528, 60170394, 83373044, 55812038, 264558, 18108385, 21206754, 5581139, 265018, 265019, 262279002, 55811150, 264563, 264682,
struct	ATPase_associated 264107, 264448	UNCLASSIFIED	struct	ATPase_associated	kinase
Contains protein domain (PF00735) - struct					
1416 94325977 (2831, 2832) Novel Protein sim. GBank gil5106557[gb AAD39749.1 AF12305 - (AF123052) MLL septin-like fusion protein [Homo sapiens]	87826663 (2833, 2834) Novel Protein sim. GBank gild 958935 dbj BAA78095.1 - (AB027570) suppressor of potassium transport defect 3 [Rattus norvegicus]		87757168 (2837, 2838) Novel Protein sim. GBank gi 2072294 (U95097) - mitotic phosphoprotein 43 [Xenopus laevis]	87298628 (2839, 2840) Novel Protein sim. GBank gi[5174421 ref NP_006023.1 pCPNE - copine VI (neuronal)	94746986 (2841, 2842) Novel Protein sim. GBank gij3876090 emb CAA93459.1 - (269635) Similarity to Yeast uridine kinase (SW:URK1_YEAST); cDNA EST EMBL:Z14695 comes from this gene; cDNA EST CEMSE17F comes from this gene; cDNA EST EMBL:067355 comes from this gene; cDNA EST yk209h1.5 comes from this gen
94325977 (2831, 2832)	87826663 (2833, 2834)	87594276 (2835, 2836)	87757168 (2837, 2838)	87298628 (2839, 2840)	94746986 (2841, 2842)
		1418	1419	1420	1421

					in the second se
	88178777 (2843, 2844)	88178777 (2843, 2844) Novel Protein sim. GBank gil4505939 ref NP_000928.1 pPOLR - polymerase (RNA) II (DNA directed) polypeptide A (220kD)		rnapolymerase	56994075, 35696286, 87168559, 55811957, 55811576, 264555, 264557, 87168518
1423				ASSIFIED	264686, 264489, 264692, 264594, 264603, 265018, 264908
1424	95201610 (2847, 2848)	95201610 (2847, 2848) Novel Protein sim. GBank gi 437181 (U02289) - GTPase- activating protein [Caenorhabditis elegans]	Contains protein domain (PF00620) - struct RhoGAP domain		29331822, 29331825, 29331827, 29146498, 264905, 264906, 264908, 264909, 265007, 264910, 265000, 3108954, 265010, 87168559, 265010, 264766, 24687, 21906765, 21906765, 21906765, 21906762, 21906766, 21906767, 21906768, 29148629, 265021, 264691, 264692, 56526486, 22279002, 264563
1425		Novel Protein sim. GBank gi 100798 pir S14959 - proline- rich protein - wheat		UNCLASSIFIED	265007, 264558
1426		94322115 (2851, 2852) Nover Protein sim. GBank gi 2078441 (U56964) - weak similarity to S. cerevisiae intracellular protein transport protein US)1 (SP:P25386) [Caenorhabditis elegans]		UNCLASSIFIED	264488, 60424179, 35696286, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 29331825, 29331827, 56182435, 264910, 6043336, 6043348, 21906766, 265018, 264288, 21906765, 21906766, 2306770, 21906769, 265020, 265022, 33657109, 18108370, 18108376, 264558, 83373044, 18108385, 56526486, 222799002, 264482
1427	91227510 (2853, 2854)	91227510 (2853, 2854) Novel Protein sim. GBank gij5616074 gb AAD45616.1 AF06194 - (AF061943) protate- derived STE20-like kinase PSK [Homo sapiens]	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	kinase	22278994, 56994075, 22278997, 29331828, 29331830, 264828, 265006, 265007, 265008, 265009, 264683, 264288, 18108354, 21906768, 29148629, 33657023, 18108374, 35695855, 83373044, 22279002, 264564
1428		94323008 (2855, 2856) Novel Protein sim. GBank gil138350 sp P28968 VGLX_HSVEB - GLYCOPROTEIN X PRECURSOR		glycoprotein	56181686, 264259, 264907, 265007, 265009, 264595, 265010, 264686, 65274620, 264629, 65274791, 22279002, 264566
1429	_			UNCLASSIFIED	264112, 264595, 265017, 265019, 21906765, 263977, 264555
1430	94735021 (2859, 2860) Novel Protein sim. Gi (D82364) a variant of	Novel Protein sim. GBank gil1181619[dbj[BAA11565] - (D82364) a variant of TSC-22 [Gallus gallus]			264094, 29331824, 264591, 264593, 265018, 264681, 21906765, 21906767, 65274620, 55811576, 264639, 87168518, 22279002
1431		80429081 (2861, 2862) Novel Protein sim. GBank gi 5420389 emb CAB46680.1 - (AJ243460) proteophosphoglyçan [Leishmania major]		UNCLASSIFIED	264768, 265020, 264691, 264693, 264638
1432		87463004 (2863, 2864) Novel Protein sim. GBank giļ414797 (L18966) - pyruvate dehydrogenase phosphatase [Bos taurus]		phosphatase	18108394, 29146498, 265007, 60433438, 264763, 29148629, 263969
1433		87605403 (2865, 2866) Novel Protein sim. GBank gil2460316 (AF022147) - uterus- ovary specific putative transmembrane protein [Rattus norvegicus]	Contains protein domain (PF00100) - UNCLASSIFIED Zona pellucida-like domain	UNCLASSIFIED	264259, 264510, 264591, 264603, 264565
1434 434	1434 85713730 (2867, 2868)	Н		UNCLASSIFIED	264682, 264691

22278997, 22278998, 22278999, 60432049, 264259, 29331824, 29331824, 66714117, 29331826, 33655970, 264508, 264905, 66712502, 29331830, 264909, 265007, 265008, 264910, 265009, 60433356, 60433438, 264596, 21906754, 265010, 265017, 265018, 265019, 18108351, 264762, 264448, 224288, 26519, 18108351, 264627, 21906768, 21906767, 21906768, 21906768, 265020, 265021, 265022, 264691, 36511576, 264636, 60170394, 56182323, 264559, 83373044, 87168518, 6042713, 22279000, 22279002, 264687, 264656, 26502, 264691, 36511576, 264636, 60170394, 56182323, 264559, 83373044, 87168518, 60432113, 22279000, 22279002,	263978, 264557, 264559	22278997, 66714117, 29331826, 264907, 56182435, 265009, 18108351, 264692, 264693	264488, 264508, 264906, 264909, 264757, 264600, 264601, 264605, 264769, 264600 3566433, 364568, 264653, 364669	2278994, 22278996, 35695286, 22278997, 22278994, 22278999, 264259, 60432049, 22278997, 22278999, 264259, 60432049, 29331822, 29331824, 60432289, 26331826, 29331827, 29331827, 29331827, 29331827, 29331827, 264509, 264509, 264907, 66712502, 29331830, 56182435, 264511, 265006, 265007, 265009, 6043229, 60433438, 264955, 55812038, 55811386, 265011, 265017, 265019, 18108354, 264448, 18108354, 264288, 18108355, 264767, 21906765, 21906766, 21906766, 21906766, 21906766, 21906767, 21906769, 55811957, 35695917, 265020, 265021, 265022, 33657109, 18108370, 18108374, 55810764, 35695655, 264634, 264887, 2643617, 264006, 264486, 264887, 264006, 264006, 264486, 264006, 2	264867, 264239, 264906, 264907, 264908, 264909, 265009, 265010, 265011, 265019, 265019, 265019, 265020, 265020, 33657023, 33657109, 60431528, 55811576, 264635
transcriptfactor	UNCLASSIFIED	UNCLASSIFIED	kinase	cathepsin	ONCLASSIFIED
			Contains protein domain (PF00595) - kinase PDZ domain (Also known as DHR or GLGF).	Serine carboxypeptidase Serine Landoxypeptidase	
1435 94708213 (2869, 2870) Novel Protein sim. GBank gij3970850 dbj BAA34789.1 - (AB015330) HRIHFB2007 [Homo sapiens]	86635024 (2871, 2872) Novel Protein sim. GBank gij3183977 emb CAA39515 - (X56044) protein Htf9C (Mus musculus)	B7631082 (2673, 2874) Novel Protein sim. GBank gi 2496867 sp Q09232 YQ22_CAEEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III	85544280 (2875, 2876) Novel Protein sim. GBank gil1905906 (AD000092) - hypothetical human serine-threonine protein kinase R31240_1 [Homo sapiens]	1438 91231894 (2877, 2878) Novel Protein sim. GBank gij3876299[emb]CAA94892 - (Z7180) similar to BPTI/KUNITZ inhibitor domain; cDNA EST EMBL:D68293 comes from this gene; cDNA EST yk448h4.5 comes from this gene; cDNA EST yk249e6.5 comes from this gene; cDNA EST yk448h4.3 comes from this gene [Caenorhabdi	(AB007902) HH0712 cDNA clone for KIAA0442 has a 574-bp insertion at position 1474 of the sequence of KIAA0442. [Homo sapiens]
5 94708213 (2869, 2870)		87631082 (2873, 2874)	8 85544280 (2875, 2876)	81231894 (2877, 2878)	-
1435	1436	1437	1438	1440	

144	1 95317662 (2881, 2882	1441 [95317662 (2881, 2882) Novel Protein sim. GBank gil4493956 jembiC4811103 21	Contract of the desired or the contract of			
		(Z98551) predicted using hexExon, MAL3P6.28	F-box domain.	- nelicase	18108392, 264488, 263994, 264489,	
		(PFC0845c), Hypothetical protein, Ien: 167 aa; Similarity to			34 35 35 35 35 35 35 35	
		model organism hypothetical proteins (C.elegans,			35036260, 22276397, 22276399, 60432049, 364360, 36534637, 36534634, 554554	
		D.melanogaster, S.cerevisiae & S.pombe). C.elegans			66714117 20221826 20331824, 26182181,	
		protein ZK287.5 (TR:			36606063 20146400 264500 201001	
_					33030032, 23 40436, 204308, 204905,	
					204503, 204806, 204507, 204508, 264909,	
_					36162453, 264310, 264311, 264312, 264910, 364603, 364603, 364603	
				_	204332, 204333, 3363/402, 60433438,	
_					204090, 204708, 21906754, 85658542,	
					67 105474, 2550110, 87168559, 264600,	
					264602, 265017, 264604, 265018, 264605,	
_					205019, 264760, 264761, 264762, 264681,	
					264448, 264764, 264683, 264288, 264766,	
					264768, 264769, 52644229, 264689,	
					21906765, 21906766, 21906767, 21906768,	
					21906769, 55811957, 35695917, 265020	
					265021, 60170615, 52644150, 264691	
					264692 33657023 65274620 32657100	
					25555502 2027023, 03274020, 33037109,	
					35595753, 254528, 18108370, 254529,	
					[18108374, 55811576, 35696423, 65274791,	_
_					35695855, 264631, 264634, 264635, 264636	
					264637 264638 56182323 264558	_
					F0170304 264630 264660 62272044	_
_					4400004, 204009, 204008, 0007,0044,	_
					16106365, 18108388, 56526486, 22279000,	
					22279002, 264563, 264483, 264564, 264566,	_
1442		83367491 (2883, 2884) Novel Protein sim. GBank gil5103027IdhilBaa78765 11			264567, 264486	
		(AB023419) mSox7 (Mus musculus)		transcriptfactor	264906, 265007, 264693, 264558	_
1443	_					_
	_	Sittoodassia siii. Obalik	Contains protein domain (PF00435) -	struct	52645080, 264691, 264628, 264555	T
		gi 486/229 gb AAD32244.1 AF15075 - (AF150755)	Spectrin repeat			
1444	2000 F000, 0710C370	microtubule-actin crosslinking factor [Mus musculus]				_
-	01050410 (2001, 2008)	Novel Prolein sim. GBank gij3874447 emb CAB02772 -		UNCLASSIFIED	264259 29331822 29331824 EE714117	7
		(281039) predicted using Genefinder; cDNA EST			20331838 264288 25505047 22552222	_
		EMBL:T01209 comes from this gene; cDNA EST			2001020, 204200, 30080817, 33037023,	_
		yk278a11.3 comes from this gene; cDNA EST vk278a11.5			204033, 00170394	_
		comes from this gene; cDNA EST vk308a9.3 comes from				_
		this gene, cDNA EST vk308a9.5 com				
1445		94990470 (2889, 2890) Novel Protein sim. GBank nil2959886lambil a 4110251				-
		(AJ222968) L-periaxin [Mus musculus]		UNCLASSIFIED	264369	
1446	85079268 (2891, 2892)	85079268 (2891, 2892)				_
1447	86945392 (2893, 2894)	Novel Protein sim GBank			264369	_
		01508161010b1440130464 11AC12C44 (AF12C440)	Contains protein domain (PF01846) -		18108396, 35696286, 22278997, 66714117,	_
_		bintington wasst product of Man American Control (Control of Control of Contr	rr domain		29331828, 265009, 264758, 265018, 264288,	-
		indicate parties of livins musculus			21906766, 21906767, 264692, 264634,	
1448		94990477 (2895, 2896) Novel Protein sim. GBank gill3980411 /AC0045611, putation protein description	Con condition		264566	_
		proline-rich protein [Arabidopsis thaliana]	Contains protein domain (PF00439) -		29331827, 264509, 264909, 265008, 264595,	_
1		Total Control of the	Diomoniali		18108357, 18108385, 264566, 264486	—

1449	1449 87860859 (2897, 2898)			UNCLASSIFIED	66714117, 264906, 264908, 264591, 264601,
1450		87458696 (2899, 2900) Novel Protein sim. GBank gij1707074 (U80450) - M01E11.2 [Caenorhabditis elegans]		UNCLASSIFIED	35696286, 35696052, 265008, 265009, 60170831, 33109954, 264683, 264689, 35696423, 35695855, 56526486
1451		Novel Protein sim. GBank gi 4160304 emb CAA10600 - (AJ132192) HS1 binding protein 3 [Mus musculus]		UNCLASSIFIED	29331826, 264683, 264693, 263978, 264630
1452		85692899 (2903, 2904) Novel Protein sim. GBank gi 2832906 dbj BAA24608.1 - (D89340) dipeptidyl peptidase III [Rattus norvegicus]		peptidase	264681, 33657023, 264629
1453		86130434 (2905, 2906) Novel Protein sim. GBank gij728831 sp P39188 ALU1_HUMAN - !!!! ALU SUBFAMILY J WARNING ENTRY !!!!		kinase	264510, 264768
1454	11204696 (2907, 2908)	_			264556
1455	1455 87797896 (2909, 2910)			UNCLASSIFIED	29331822, 66714117, 29331825, 264905, 29331830, 265006, 265008, 265009, 265011, 265019, 18108351, 21906768, 33657109, 18108376, 264832, 6418333, 87168518
1456		86320218 (2911, 2912) Novel Protein sim. GBank gij729230 sp P41004 CUT3_SCHPO - CHROMOSOME SEGREGATION PROTEIN CUT3		transport	22278995, 22278986, 22278997, 22278986, 22278999, 22278999, 29331827, 264107, 265017, 21906765, 21906766, 21906767, 21906769, 18108370, 222789000
1457	80076900 (2913, 2914)	80076900 (2913, 2914)		UNCLASSIFIED	264107, 264566
1458	87800460 (2915, 2916)	Novel Protein sim GBank dil 2246532 (193872) - ORE 73		tronsport	E6190676 22278000 60422040 264260
		contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		Todsum	201 82279, 22278999, 50442049, 264259, 26331826, 26331827, 29331828, 264102, 264107, 265009, 60432229, 265019, 265020, 263976, 264635, 22279002, 264566
1459	95360920 (2917, 2918)	95360920 (2917, 2918) Novel Protein sim. GBank gil5524667[gb AAD44333.1 AF15935 - (AF159356) Munc13- C2 domain 4 protein [Rattus norvegicus]	Contains protein domain (PF00168) - kinase C2 domain		22278997, 284259, 25331824, 29331826, 29331827, 29331828, 265017, 265018, 264760, 264682, 264448, 264288, 264766, 265021, 264692, 33657023, 33657109, 35695855, 264566
1460				UNCLASSIFIED	29331822, 264591, 55811957, 264691, 264693, 65274620
1461		94741513 (2921, 2922) Novel Protein sim. GBank gi 1707274 (U80931) - strong similarity to class-III of pyridoxal-phoshate-dependent aminotransferases [Caenorhabditis elegans]	Contains protein domain (PF00202) - gaba Aminotransferases class-III pyridoxal phosphate	gaba	22278997, 29331822, 35696052, 265009, 264750, 264750, 264369, 264607, 265018, 265019, 264750, 264369, 264687, 21906765, 21906768, 265022, 33657109, 27486261, 264555, 83373044
1462				UNCLASSIFIED	264555, 264556
1463	88090605 (2925, 2926) Novel Protein sim. ((X98259) M-phase	Novel Protein sim. GBank gi 1770466 emb CAA66912 - (X98259) M-phase phosphoprotein 8 [Homo sapiens]	Contains protein domain (PF00385) - struct 'chromo' (CHRromatin Organization	struct	60432049, 264259, 29146499, 264906, 264907, 264512, 265017, 264763, 264766,
			MOdifier) domain		18108370, 18108374, 264636, 18108385, 18108 <u>38</u> 8

264569, 22278995, 22278996, 22278997, 22278998, 29331822, 29331824, 29331825, 35696052, 20281100, 264905, 29331830, 264909, 265007, 33657402, 21906754, 265017, 265018, 264682, 264368, 264766, 21906766, 21906766, 21906766, 21906766, 21906769, 33657023, 264692, 264693, 264692, 2646931, 264631, 264633, 2	264488, 22278994, 56994075, 60432049, 264259, 56182181, 60432289, 29331827, 5264045, 264511, 265007, 265008, 264596, 55812038, 55811386, 264600, 264602, 265007, 265018, 265014, 265019, 18108351, 18108354, 56181562, 21906769, 265021, 18108352, 60432113,	264512, 265017, 264689, 264558	264683, 264636	22278998, 264758, 265018, 265019, 21906769, 265020, 33657109, 22279002	18108394, 18108397, 18108398, 35696052, 29146499, 265007, 265008, 265009, 265010, 265011, 18108354, 18108365, 18108388, 18108374, 18108381, 18108382, 18108388, 18108388	18108394, 18108398, 56182575, 264259. 29331822, 29331824, 29331825, 60432289, 264907, 264909, 265007, 264910, 265009. 264591, 6043229, 60433356, 264595, 60433438, 264758, 33109954, 265010, 265011, 265018, 264760, 26448, 264764, 265021, 265018, 264760, 26448, 264764, 265021, 18108361, 1810837, 264769, 56811957, 265021, 18108361, 264691, 18108362, 18108362, 18108381, 56182323, 18108382, 83373044, 18108384, 18108388, 87168518, 60432113, 264404, 22279002, 264482, 264567, 264482, 264697, 264482,	29331822, 29331824, 29331825, 264828, 264603, 264689, 264693, 18108374, 55811576
UNCLASSIFIED	glucoamylase	UNCLASSIFIED	struct		struct	UNCLASSIFIED	transport
							Contains protein domain (PF00153) - Iransport Mitochondrial carrier proteins
1464 87620482 (2927, 2928) Novel Protein sim. GBank gij3874447 emb CAB02772 - (Z81039) predicted using Genefinder; cDNA EST EMBL: T01209 comes from this gene; cDNA EST yk278a11.3 comes from this gene; cDNA EST yk278a11.5 comes from this gene; cDNA EST yk378a11.5 comes from this gene; cDNA EST yk308a9.3 comes from this gene; cDNA EST yk308a9.5 com	87425192 (2929, 2930) Novel Protein sim. GBank gilę589598 dbj BAA76821.1 - (AB023194) KIAA0977 protein [Homo sapiens]		87614328 (2933, 2934) Novel Protein sim. GBank gi[2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]	1468 95342862 (2935, 2936) Novel Protein sim. GBank gil4507241 ref NP_003137.1 pSSRP - structure specific recognition protein 1	Novel Protein sim. GBank giļ1906596 (U81788) - kinesin-73 [Drosophila melanogaster]	94890482 (2939, 2940) Novel Protein sim. GBank gi 5549170 gb AAD43131.2 AF15909 - (AF159092) syld709613 protein [Homo sapiens]	87826842 (2941, 2942) Novel Protein sim. GBank gij3875146jemb CAB01750] - (Z78542) similar to Mitochondrial carrier proteins; cDNA EST EMBL:T01651 comes from this gene [Caenorhabditis elegans]
87620482 (2927, 2928) (87425192 (2929, 2930)	87606227 (2931, 2932)	87614328 (2933, 2934) I	95342862 (2935, 2936) I	79236174 (2937, 2938) 1		87826842 (2941, 2942) (
1464	1465	1466	1467	1468	1469	1470	1471

1472	87756616 (2943, 2944)	1472 87756616 (2943, 2944) Novel Protein sim. GBank gil4680707[gblAAD27743.1 AF13296 - (AF132968) CGI-34 protein [Homo sapiens]		UNCLASSIFIED	264905, 18108351, 21906765, 264486
1473		87791609 (2945, 2946) Novel Protein sim. GBank gij3688780 (AF042180) - testis-specific Y-encoded-like protein [Mus musculus]	Contains protein domain (PF00956) - MHC Nucleosome assembly protein (NAP)	MHC	18108394, 22278995, 56994075, 22278999, 29331822, 29331824, 66714117, 29331825, 29331826, 35696052, 264906, 264907, 265182435, 256007, 264758, 255018, 255019, 264760, 264764, 264289, 264685, 264686, 264768, 21906769, 55811957, 265021, 264631, 264632, 55811576, 264634, 264638, 26182323, 22279002, 264566, 264886
1474	85800989 (2947, 2948)	1474 85800989 (2947, 2948) Novel Protein sim. GBank gi[2494890]sp[092176]CORO_BOVIN - CORONIN-LIKE PROTEIN P57	Contains protein domain (PF00400) - struct WD domain, G-beta repeat	struci	264488, 35695917, 35696286, 264692, 33657023, 264693, 33657109, 35696052, 264508, 264905, 284907, 264629, 264508, 264909, 35696423, 35693855, 264511, 264910, 264632, 264634, 264635, 264536, 264537, 264556, 264557, 264639, 264556, 2647241, 264604, 264605, 264566,
1475	86871935 (2949, 2950)		Contains protein domain (PF00041) - UNCLASSIFIED Fibronectin type III domain	UNCLASSIFIED	264681, 264682, 264288, 264566
1476	87548655 (2951, 2952)	87548555 (2951, 2952) Novel Protein sim. GBank gi 4757752 ref NP_004654.1 pANGP - angiopoietin 3	Contains protein domain (PF00147) - glycoprotein Fibrinogen beta and gamma chains, C-terminal globular domain	glycoprotein	60424179, 56181686, 29331824, 60424269, 29331826, 35696052, 264508, 264905, 264907, 264908, 264908, 264905, 265007, 265008, 265009, 264910, 33657402, 264595, 264596, 55812038, 265011, 264501, 264762, 18108351, 264288, 264369, 264989, 264699, 264693, 18108370, 60431528, 18108374, 35696423, 264634, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264638, 264636, 264636, 264638, 264638, 264636, 264638, 2646
1477	87774279 (2953, 2954)	87774279 (2953, 2954) Novel Protein sim. GBank gi Z498308 sp Q60870 DP1_MOUSE - POLYPOSIS LOCUS PROTEIN 1 HOMOLOG (TB2 PROTEIN HOMOLOG) (GP106)		UNCLASSIFIED	264259, 264107, 264905, 265008, 265010, 265011, 264682, 264288, 265020, 265021, 263974
1478	1478 11754412 (2955, 2956)				264686

56182575, 22278995, 22278996, 22278998, 22278999, 29331822, 29331824, 66714117, 264906, 264907, 56182435, 265006, 60170831, 33657402, 264758, 33109954, 21906754, 265017, 265019, 264448, 264286, 264767, 264687, 5264429, 21906764, 265020, 265021, 60170615, 264691, 2365702, 33657149, 18108370, 60431528, 263976, 55811576, 264566, 264557, 60170394, 87168518, 264404, 22279000, 22279002, 264563, 264482	16108394, 62274572, 56182575, 22278995, 35696286, 56994075, 22278996, 22278997, 22278999, 22478999, 22278999, 22278999, 22278999, 22278999, 22478999, 22478999, 2244091, 264405, 264108, 265009, 60433366, 60433438, 264596, 33109954, 33657084, 52644296, 87168474, 265019, 87168559, 265017, 265019, 264448, 265020, 265021, 266769, 21906766, 21906766, 21906765, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 2650201, 266972, 35692855, 264577, 2569885, 264557, 263981, 264089, 264486	265006, 265007, 265010, 18108374 264908, 264910, 264758	264259, 29331822, 52645080, 29331825, 29331826, 33656970, 29331830, 266007, 55812038, 33109954, 265017, 264288, 21906768, 21906769, 264636, 18108380, 87168518, 22279000	264489, 35696286, 264259, 264107, 264909, 265008, 60433356, 33657402, 60433438, 264288, 21906765, 21906766, 29148627, 33657023, 27486262, 18108374, 35696423, 83373044, 60432113	264369, 265020, 18108374
peptidase	UNCLASSIFIED	MHC UNCLASSIFIED	struct	UNCLASSIFIED	- struct
		MHC Contains protein domain (PF00008) - UNCLASSIFIED	EGF-like domain Contains protein domain (PF00036) - EF hand		Contains protein domain (PF00225) - struct Kinesin motor domain
Novel Protein sim. GBank gi[5499741]gb[AAD43978.1]AF15296 - (AF152961) chromatin-specific transcription elongation factor FACT 140 kDa subunit [Homo sapiens] kDa subunit [Homo sapiens]	Novel Protein sim. GBank gij3560456 emb CAA06329.1 - (AJ005073) Alix [Mus musculus]	87021442 (2961, 2962) Novel Protein sim. GBank gil48368071gb AAD30566.1 AF14679 - (AF146793) PFT27 [Mus musculus] 85320442 (2963, 2964) Novel Protein sim. GBank	gil4585372lgb AAD25403.1/AF12292 - (AF122923) Wnt EGF-like domain inhibitory factor-1 [Mus musculus] 94115503 (2965, 2966) Novel Protein sim. GBank gil535428 (U13738) - calmodulin- Contains protein domain (PF00036) - struct [ike protein [Pisum sativum]	Novel Protein sim. GBank gil1911774 bbs 180090 - (S83364) putative Rab5-interacting protein (clone L1-57) [human, HeLa cells, Peptide Partial, 122 aa] [Homo sapiens]	1485 80194441 (2969, 2970) Novel Protein sim. GBank gil5360129[gb]AAD42883.1 AF15511 - (AF155117) NY-REN Kinesin motor domain 62 antigen [Homo sapiens]
1479 91640140 (2957, 2958) Novel Protein sim. GBank gi[5499741]gb[AAD43978. chromatin-specific transcri kDa subunit [Homo sapier	94312412 (2959, 2960) Novel Protein sim. G (A_J005073) Alix [Mu	$\overline{}$		4 94131544 (2967, 2968) Novel Protein sim. (S83364) putative I [human, HeLa cells sapiens]	5 8019441 (2969, 2970)
1478	1480	1481	1483	1484	4

14RG	14RR 04125066 (2071 2072) Mousi Destein sim	Mondi Bratain aim Charle allegaceachar in a annan a			
<u>}</u>	101(2000 (20) 1, 20(2)	Mover Protein Sun. Grank gil4569516 apil6AA/6/80.11	Contains protein domain (PF00069) - kinase	kinase	56182575, 22278999, 264906, 264907,
	_	(Abuta 133) NiAAU330 protein [Homo sapiens]	Eukaryotic protein kinase domain		21906754, 87168474, 265017, 265019,
					18108351, 264288, 265020, 264566
148/	(86452711 (2973, 2974) Novel Protein sim.	Novel Protein sim. GBank gij5019275/emb CAB44431.1 -		synthase	21906754, 264486
		(AJ132751) xenobiotic/medium-chain fatty acid:CoA ligase			
17.0		rorm AL-III Bos taurus			
- - - -	0//32026 (29/5, 29/6) Novel Protein sim.	Novel Protein sim. GBank	Contains protein domain (PF01443) - fgf	fgf	264686, 264769, 264689, 264692, 264693,
_		gi[5712131]gb AAD47379.1 AF12049 - (AF120499) DEM1	Viral (Superfamily 1) RNA helicase		264509, 264906, 264907, 18108370, 264908,
		protein [Homo sapiens]			264629, 264909, 264510, 265006, 264512,
					265007, 265008, 265009, 264555, 264556,
400	100 CT00/ CT070F0				264557, 264558, 264762, 264564, 264682
20	1408 93104277 (2877, 2978) Novel Protein Sim.		Contains protein domain (PF00047) - prostaglandin	prostaglandin	21906767, 22278999, 265022, 264259,
		gilz49/303jsp Q62786 FPRP_RAT - PROSTAGLANDIN F2- Immunoglobulin domain	:- Immunoglabulin domain		264693, 29331824, 29331825, 29331826,
		ALPHA RECEPTOR REGULATORY PROTEIN			29331827, 29331828, 264103, 263972,
		PRECURSOR (PROSTAGLANDIN F2-ALPHA RECEPTOR			66712502, 35696423, 35695855, 265007,
		ASSOCIATED PROTEIN)			265008, 265009, 83373044, 21906754,
					56526486, 265017, 264563, 18108351,
40	87300177 2000				264564, 264566, 264369, 264288
1	1430 01330127 (2373, 2360)			UNCLASSIFIED	56182575, 264259, 29331822, 29331824,
					66714117, 29331827, 29331828, 264508,
					264905, 66712502, 265007, 265008, 264594
					33657402 55812038 87168474 265018
_					40400014 00414000 01100414, 200010,
	_				10100331, 204309, 204266, 204709, 204089,
					21906767, 21906768, 55811957, 60170615,
					33657109, 35695855, 264635, 60170394,
3	Т	- 1			56526486, 22279002, 264563
148	83394305 (2981, 2982) Novel Protein sim.	Novel Protein sim. GBank gij295671 (L11275) - selected as		UNCLASSIFIED	265007, 264448, 18108372, 264558
		a weak suppressor of a mutant of the subunit AC40 of DNA			56182323
	-	dependant RNA polymerase Land III (Saccharomyces			
1492	1492 [85805363 (2983 2984) Novel Protein sim	-	Contract and an ariabase		
	(100 100	musculus]	Contains protein domain (PF00071) - oncogene Ras family	oncogene	22278997, 22278998, 29331822, 264907, 66712603
					2007 100

264488, 52646365, 65274572, 56182575, 22278994, 35696286, 56994075, 22278999, 56432049, 35696286, 56994075, 22278999, 56432049, 256908, 56994075, 22278999, 56432049, 264905, 264908, 2564008, 265009, 266182455, 265006, 265008, 265000, 266017, 265017, 265018, 265019, 55811150, 265011, 265017, 265018, 265019, 55811150, 265010, 265017, 265018, 21906768, 21906768, 21906769, 21906765, 21906768, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 21908362, 27486262, 27486262, 27486262, 27486262, 27486262, 27486264, 2648626, 21908370, 264129, 18108374, 52644332, 56182323, 87168518, 22279002, 264564, 264566, 264567	264907, 255009, 264759, 18108370, 55811576, 264639, 264565, 264486	22278997, 264259, 264906, 264907, 265009, 264594, 33657084, 265017, 264760, 264448, 33657109, 264630, 264634, 56526486, 264565, 264566, 264486, 264567	4906, 265019, 264693	264910, 264758, 265011, 264764, 264288, 264690, 264634, 264635, 56526486	264488, 65274572, 29331822, 265017, 265018, 21906765, 29148627, 29148629, 18108374, 264637, 264638, 264567	999	22278999, 264769, 18108379		264508, 264112, 264604, 264684, 52644150, 55811576, 264632, 264556, 264638, 56182323, 264563, 264486	29331822, 265007, 264369
264488, 5264 22278994, 35 60432049, 29 264508, 2649 56182435, 26 60170817, 2650 265011, 2650 265020, 2650 265020, 2650 3365723, 18 33657182, 27 264629, 1810 87168518, 22	264907, 2650 55811576, 26	22278997, 26 264594, 3365 33657109, 26 264563, 2645	35696286, 28	264910, 264; 264690, 264	264488, 652 265018, 219(18108374, 26	263978, 264566	22278999, 20	264559	264508, 264 55811576, 24 56182323, 20	29331822, 2
UNCLASSIFIED	kinase	transcriptfactor	UNCLASSIFIED	kinase	oucogene	glucoamylase	UNCLASSIFIED	ÜNCLASSIFIED		transferase
	Contains protein domain (PF01352) - kinase KRAB box	Contains protein domain (PF01352) - Iranscriptfactor KRAB box		Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain						Contains protein domain (PF00535) - Iransferase Glycosyl transferases
91677215 (2985, 2986) Novel Protein sim. GBank gil5689515 dbj BAA83041.1 - (AB029012) KIAA1089 protein [Homo sapiens]	87605265 (2987, 2988) Novel Protein sim. GBank gij728832lspjP39189JALU2_HUMAN - IIII ALU SUBFAMILY KRAB box SB WARNING ENTRY IIII	87605267 (2989, 2990) Novel Protein sim. GBank gil4589588 dbj BAA76816.1 - (AB023189) KIAA0972 protein [Homo sapiens]	87784322 (2991, 2992) Novel Protein sim. GBank gi[5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]	81695428 (2993, 2994) Novel Protein sim. GBank gi[3874925 emb CAA92591 - (268296) Similarity to Mouse A-RAF proto-oncogene serine/threonlhe-protein kinase (SW:KRAA_MOUSE); cDNA EST EMBL:D27610 comes from this gene; cDNA EST EMBL:T01018 comes from this gene; cDNA EST EMBL:D33256 comes from this gene; cDNA EST	1498 90934938 (2995, 2996) Novel Protein sim. GBank gij728836 sp P39193 ALU6_HUMAN - !!!! ALU SUBFAMILY SP WARNING ENTRY !!!!	86451589 (2997, 2998) Novel Protein sim. GBank gil2570198 (U54556) - microfitarial sheath protein SHP3 [Litomosoides sigmodontis]		85795297 (3001, 3002) Novel Protein sim. GBank gij 2078483 (U43200) - antifreeze glycopeptide AFGP polyprotein precursor [Boreogadus saida]		87012701 (3005, 3006) Novel Protein sim. GBank gil3900855 (AC004874) - similar to N-acetylgalactosaminyttransferase; similar to Q07537 (PID:g1171989) [Homo sapiens]
91677215 (2985, 2986) P	87605265 (2987, 2988)	87605267 (2989, 2990)		81695428 (2993, 2994)	90934938 (2995, 2996)	86451589 (2997, 2998)	80499386 (2999, 3000)	85795297 (3001, 3002)	80206141 (3003, 3004)	87012701 (3005, 3006)
1493	1494	1495	1496	1497	1498	1499	1500	1501	1502	1503

Contains protein domain (PF00023) - UNCLASSIFIED
Contains protein domain (PF00069) - Eukaryotic protein kinase domain
Contains protein domain (PF00018) - glycoprotein SH3 domain
Novel Protein sim. GBank Contains protein domain (PF00069) - kinase gil5031975/reflNP_005875.1jpPAK4 - protein kinase related Eukaryotic protein kinase domain to S. cerevisiae STE20, effector for Cdc42Hs
Novel Protein sim. GBank gil4559353igb AAD23014.1 AC00658 - (AC006585) putative RIO1/ZK632.3/MJ0444 family extragenic suppressor protein [Arabidopsis thaliana]

52644507, 52645156, 52646366, 52646842, 65274572, 22278994, 35696286, 56994075, 264259, 52645080, 29331822, 29331822, 29331822, 29331822, 29331823, 25686052, 29331830, 52644045, 56182435, 265006, 60433356, 60433438, 55812038, 21906754, 52644299, 21906765, 21906765, 21906766, 21906767, 21906768, 35695917, 265020, 52644150, 33657023, 52645129, 33657109, 33657182, 27486261, 27486262, 27486264, 35695855, 52644323, 18108336, 35696423, 35695855, 5264432, 18108336, 35696423, 35695855, 5264432, 18108385, 35696423, 35695855, 52644322, 18108385, 35696423, 35695855, 52644322, 18108385, 35696423, 35695855, 52644322, 18108385, 35696423, 35695855, 52644322, 18108385, 35696423, 35695855, 52644322, 35696423, 35695855, 52644322, 35696423, 356964322, 356964324, 35696423, 35696423, 35696423, 35696423, 356964324, 356964324, 356964324, 356964324, 3569644324, 3569644324, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964434, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964424, 356964444, 356964444, 35696444, 35696444, 35696444, 35696444, 35696444, 35696444,	16106367, 87108318, 60432113 265020, 264639	265008, 56182323, 22279002 264091, 18108370, 264404	66714117, 284508, 264509, 264905, 264510, 264910, 264511, 264591, 264595, 264288, 264786, 264769, 18108374, 264636, 264638, 264486	264569, 264489, 60432049, 265009, 33657402, 264596, 21906754, 265019, 264369, 21906765, 21906768, 21906769, 264691, 65274620, 33657182, 27486261, 18108374, 264557, 264639, 87168518.	18108392, 60432049, 264259, 29331824, 265007, 60433356, 265010, 21906768, 264636	18108394, 264259, 66714117, 265011, 264603, 265019, 18108364, 35696423, 264557, 264558, 18108388	29331824, 265018, 265020, 265021 65274572, 21906768, 264693	264112, 21906754, 263974 264683, 264687, 264689, 264692, 264693
- UNCLASSIFIED	ubiquitin	UNCLASSIFIED	tm7		UNCLASSIFIED	synthase	UNCLASSIFIED	UNCLASSIFIED UNCLASSIFIED
Contains protein domain (PF01163) - UNCLASSIFIED RIO1/ZK632.3/MJ0444 family						Contains protein domain (PF00483) - synthase Nucleotidyl transferase		
I = 10 10	79163536 (3029, 3030) Novel Protein sim. GBank gij3879501[emb[CAA87795] - (Z47812) similar to ubiquitin carboxyl-terminal hydrolase; cDNA EST EMBL:D33366 comes from this gene; cDNA EST EMBL:D33965 comes from this gene; cDNA EST EMBL:D33822 comes from this gene; cDNA EST EMBL:D33822 comes from this gene; cDNA EST	880 / 35J9 (3U31, 3U32) Novel Protein sim. GBank gil498015 (L27479) - X123 [Homo sapiens] 87793325 (3U33, 3U34) Novel Protein sim. GBank gil3415134 (AF082024) - Phyb1	Pimpinella brachycarpa 87350697 (3035, 3036) Novel Protein sim. GBank gil728838 sp p39195 aLU8_HUMAN - !!!! ALU SUBFAMILY SX WARNING ENTRY !!!!	Nover Protein sim. Cbank gilszözöb lembjcAB45771.11. (AL080198) hypothetical protein [Homo sapiens]	Orbazopo (3034, 3044)) Novel Protein sim. GBank gi 2662161 dbj BAA23712 - (AB007900) HH0452 cDNA clone for KIAA0440 has a 438-bp insertion at position 1711 of the sequence of KIAA0440. Monos appins)	1/AF13542 - (AF135421) GDP. se B [Homo sapiens]	91005151 (3045, 3046) Novel Protein sim. GBank gi[3776567 (AC005388) - Strong similarity to F2187.33 gi[2809264 from A. thaliana BAC gblAC002560. EST gblA65119 comes from this gene. [Arabidopsis thaliana]	80203723 (3047, 3048) 87799887 (3049, 3050) Novel Protein sim. GBank gjl4759040[ref]NP_004283 1]pRIN1 - ras inhibitor
95345392 (3027, 3028)	79163536 (3029, 3030)		87350697 (3035, 3036)	(000) (000)	07 39 20 30 39 30 40)	78960687 (3043, 3044)	91005151 (3045, 3046)	80203723 (3047, 3048) 87799867 (3049, 3050) N
 	1515	1517	1518		1521	1522	1523	1524

1526	1528 95105344 (2051 2052) Noviel Protein eim Ci	Novel Protein cleant			
		gij728850 sp P08640 AMYH_YEAST - GLUCOAMYLASE	J	alycopioses:	-55812038, 56162161, 60431735, 264595, 55812038, 264605, 264683, 21906765,
		S1/S2 PRECURSOR (GLUCAN 1,4-ALPHA-			55811957, 265020, 65274791, 264555.
	,	GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN			264556, 264557, 264558, 264559, 83373044
1527	88262512 (3053, 3054) Novel Protein sim, GE	Novel Protein sim GBank 0/12792496 (AF041107) - tulin 2			E8102575 DE4250 60422040 20224022
		[Rattus norvegicus]			501023/3, 204239, 00432049, 29331022, 604333380, 364008, 66743609, 60433438
					DOTAGEED DEFOAT DOLLD DARBOTTO
					6/166559, 255017, 254288, Z1906/56,
					21900/09, 2639//, 556115/6, 56182323,
1528	94130918 (3055, 3056)			INCLACCICIED	10100301 2272000E 2272005 2642E0 66742E0
			•		22210393, 22210391, 264239, 66/12302,
					264596, 265017, 265019, 264682, 264448,
					264683, 264764, 264685, 264686, 21906765,
					21906766, 21906767, 21906768, 21906769,
200	_				265022, 264693, 83373044, 18108385
67C1		94 (20/33 (3037, 3038) Novel Protein sim. GBank gil4406663/gb[AAD20053]		UNCLASSIFIED	264488, 263994, 56182575, 22278995,
		(AF131825) Unknown [Homo sapiens]			35696286, 22278997, 264259, 29331822,
					60432289, 29331827, 35696052, 264509,
					264906, 264907, 264908, 264909, 52644045,
					56182435, 264511, 265009, 264910,
					60433356, 60433438, 265017, 265018,
					264760, 264448, 264764, 264369, 264288,
					264766, 18108357, 264768, 52644229,
_					21906765, 21906766, 21906767, 21906768,
					265021, 265022, 52644150, 33657109,
					264629, 35695855, 60432113, 22279002,
	_				264563, 264564, 264486, 264567
1530	_	95012765 (3059, 3060) Novel Protein sim. GBank gi 2828710 (AF043642) - matrin			264488, 264489, 35696286, 29331825,
		cyclophilin [Rattus norvegicus]			35696052, 264508, 264905, 264906, 264907,
					264909, 264510, 264511, 264512, 264910,
				-	264592, 264595, 18108351, 264764, 264683,
				-	264684, 264766, 264768, 18108357, 264769,
					35695917, 264628, 264629, 18108374,
					35695855, 264630, 264631, 264634, 264555,
					264636, 264637, 264404, 264563, 264566,
		П			264486
5	95419351 (3061, 3062)	3ank gij 1905874 (U90878) - carboxyl	Contains protein domain (PF00595) - kinase		56182575, 35696286, 264097, 264259,
_		terminal LIM domain protein [Homo sapiens]	PDZ domain (Also known as DHR or		29331822, 29331825, 29331826, 29331827,
			GLGF).		35696052, 264509, 56182435, 264510,
				•	264511, 265007, 60433356, 55811386,
					264681, 264369, 264288, 264766, 264687,
					55811957, 35695917, 33657023, 35695763,
				-	55810764, 35696423, 55811576, 263981,
					60170394, 56182323, 83373044, 60432113,
					264566

29331824, 29146499, 264907, 264112, 265008, 265011, 265017, 265018, 264762, 18108351, 263967, 20281149, 18108374, 264568	264905, 264907, 264766, 264637	65274572, 22278997, 264259, 60432049, 29331622, 60432289, 29331627, 29146499, 265006, 265008, 60170831, 60433438, 33109954, 8716859, 265018, 18108357, 21905768, 29148629, 265021, 265022, 18108377, 56182323, 60432113, 22279000,	264686, 18108357, 18108394, 21906767, 21906768, 29148629, 35696286, 265020, 265021, 52644150, 264693, 66714117, 29331825, 29331826, 264909, 18108374, 35696423, 35695855, 265009, 264634, 264638, 264638, 264563, 264563, 264563, 264765, 18108385, 56526486, 265017, 265018, 264766, 264765, 18108351, 264448, 264369, 264766	65274572, 56182575, 60432049, 264259, 29331826, 265006, 265007, 60433356, 60433438, 264601, 18108351, 26448, 264369, 264288, 33657023, 65274620, 33657109, 60432113	18108394, 18108397, 264909, 265008, 265009, 265000, 18108351, 264638, 18108385, 18108388
UNCLASSIFIED	UNCLASSIFIED		transcriptfactor		kinase
			Contains protein domain (PF00096) - transcriptfactor Zinc finger, C2H2 type		Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain
Novel Protein sim. GBank gil3874716 emb CAA91265 - (266494) cDNA EST EMBL:D65271 comes from this gene; cDNA EST EMBL:D654845 comes from this gene; cDNA EST EMBL:D64480 comes from this gene; cDNA EST EMBL:D67438 comes from this gene; cDNA EST EMBL:D68087 comes from this gene; cDNA Novel Protein sim. GBank gil1490324 emb CAB01543 - (278141) unknown [Mus musculus]	Novel Protein sim. GBank gi 81286 pir S22697 - extensin - Volvox carteri (fragment)		Novel Protein sim. GBank gl 106024 pir B32891 - finger protein 2, placental - human		Novet Protein sim. GBank giļ403440 (M81787) - [Gallus domesticus skeletal muscle mRNA, parlial cds.], gene product [Gallus gallus]
1532 85718224 (3063, 3064) Novel Protein sim. (266494) cDNA EST EMBL: CDNA EST EMBL: D6444 EST EMBL: D67438 con EMBL: D6738 con EMBL: D66087 c		90936732 (3069, 3070)	87602856 (3071, 3072) Novel Protein sim. protein 2, placenta	95354556 (3073, 3074) Novel Protein sim. (Z79754) cDNA ES CDNA EST EMBL:1 EST yk426f12.5 cc yk342f10.5 comes comes from this ge	85724628 (3075, 3076) Novel Protein sim. domesticus skeleta product (Gallus gall
1532	1534				1538

nuclease 22278994, 22278996, 35696286, 56994075, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 60424269, 29331825, 260432289, 29331826, 29331826, 29331826, 29331828, 255009, 33657402, 6043335, 6043338, 265009, 33657402, 60433356, 60433336, 6043338, 25817038, 21906754, 81168559, 265017, 265018, 265019, 18108359, 21906764, 21906767, 21906768, 35695917, 265020, 265021, 21906768, 35695917, 265020, 265021, 218108376, 33657109, 27468261, 18108370, 18108376, 36595423, 55811576, 65274791, 264558, 56182323, 60170394, 83373044, 87168518, 60432113, 22279000, 22627866	UNCLASSIFIED 264369, 264691, 263978	phosphatase 264488, 264489, 22278999, 264259, 29331822, 35696052, 264508, 264509, 264905, 264907, 264908, 264501, 264512, 264910, 264592, 264761, 264761, 264761, 26478, 264761, 26478, 264781, 264781, 264781, 264781, 264781, 264781, 264781, 264781, 264781, 264781, 264781, 264781, 264781, 26481, 264781, 26481,	cathepsin 18108392, 18108394, 62274572, 29331822, 264508, 265007, 265008, 265001, 264682, 18108354, 18108355, 52644150, 18108368, 264636, 18108381, 18108382	nuclease 264106, 33109954, 265019, 264683, 35695917, 264690, 264692, 33657109	Contains protein domain (PF00439) - nucl_recpt 35696286, 56994075, 22278999, 35696052, Bromodomain 33657109, 35696423, 284631, 87168518, 22279000		UNCLASSIFIED 264905, 264686 UNCLASSIFIED 264259, 29331822, 66714117, 265007, 55811386, 263010, 264600, 265017, 265019, 264288, 264768, 265020, 265022, 55811576,
1539 95337628 (3077, 3078) Novel Protein sim. GBank gij3218411 emb CAA19575.1 - (AL023859) SPBC19C7.07c, putative RNA splicing endonuclease ga mma subunit, ten.284aa, similar eg. to YAR008W, YAHB_YEAST, P39707, YAR008W, tRNA splicing endonuclease gamma subunit. (275aa), fasta scores, opt.269, E():6.4e-2) Novel Protein sim. GBank gi 5052634 gb AAD38647.1 AF14567 - (AF145672) BcDNA GH12174 [Drosophila melanogaster]	95317948 (3081, 3082) Novel Protein sim. GBank gil5052349 gb AAD38515.1 AF13501 - (AF135016) protein phosphatase 2A 48 kDa regulatory subunit [Homo sapiens]) Novel Protein sim, GBank gi 5305702 gb AAD41779.1 AF12686 - (AF126867) calpain- like protease [Mus musculus]	84348768 (3085, 3086) Novel Protein sim. GBank gi 728832 sp P39189 ALU2_HUMAN - !!!! ALU SUBFAMILY SB WARNING ENTRY !!!!	·	85757973 (3089, 3090) Novel Protein sim. GBank gi 1086591 (U41007) - similar to S. cervisiae nuclear protein SNF2 (SP:P22082) in a a region of gly-arg repeats (Caenorhabditis elegans)	8699594 (3093, 3094) Novel Protein sim. GBank gi 2661132 (AF035683) - p21 [Mus musculus]
1539 95337628 (3077, 3078)							1547 86999594 (3093, 3094)

						_
29331824, 60431528, 264639, 56182323	60424179, 22278995, 35696286, 22278998, 22278999, 264094, 2933182, 56182181, 29331824, 35696052, 264905, 264906, 264906, 264909, 265006, 264911, 265008, 60431735, 60433356, 21906754, 55811386, 87168559, 265017, 265018, 265019, 55811150, 264682, 264288, 264369, 56181562, 264769, 21906768, 21906769, 56181562, 264769, 286916, 33657109, 60431528, 35696423, 35695656, 264568, 264566	65274572, 60432289, 265008, 264910, 265011, 265017, 265019, 264768, 56182323	22278999, 29331822, 264508, 264509, 264906, 264907, 264909, 265007, 264912, 264910, 21906754, 265018, 265019, 264681, 264764, 264766, 264688, 264769, 21906769, 264688, 264555, 264556, 264658, 264555, 264556, 264658, 264557, 264558, 264568, 264568, 264556, 264568, 264688	264259, 29331828, 66712502, 264764, 264288, 264686, 33657109, 264556	56182575, 35699286, 29146499, 264509, 264907, 264909, 56182435, 265006, 265008, 265008, 264910, 264757, 264758, 265017, 55811150, 18108351, 264764, 56181562, 35695917, 264693, 33657109, 18108374, 35695917, 264693, 33657109, 264655, 56182323, 18108382, 264655, 56182323, 18108382,	26696286, 22279699, 264259, 29331822, 29331824, 29331824, 29331825, 29331828, 265007, 265009, 60432229, 33657402, 55812038, 265011, 265019, 264681, 264369, 264686, 264767, 264768, 21906765, 21906769, 2645017, 264693, 18108370, 60431528, 55811576, 264631, 60170394, 56182323, 533373044, 18108385, 22279000, 22279000
UNCLASSIFIED	d do d	- tm7	UNCLASSIFIED	UNCLASSIFIED	dehydrogenase	transferase
		Contains protein domain (PF00001) - tm7 7 transmembrane receptor (rhodopsin family)	(fundamental modernia)		Contains protein domain (PF00304) - dehydrogenase Gamma-thionins family	Contains protein domain (PF00446) - transferase Gonadotropin-releasing hormones
1548 94233065 (3095, 3096) Novel Protein sim. GBank gij3043692 dbj BAA25510 - (AB011156) KIAA0584 protein [Homo sapiens]	95330048 (3097, 3098) Novel Protein sim. GBank gij5689519 dbj BAA83043.1 - (AB029014) KIAA1091 protein [Homo sapiens]	Novel Protein sum. GBank gil544463 sp P35350 GUSB_BOVIN - POSSIBLE GUSTATORY RECEPTOR TYPE B (PPR1 PROTEIN)	88077111 (3101, 3102) Novel Protein sim. GBank gil4758566 ref NP_004798.1 pHS6S - heparan-sulfate 6- sulfotransferase		94725512 (3105, 3106) Novel Protein sim. GBank gil4589570 dbj BAA76807.1 - (AB023180) KIAA0963 protein [Homo sapiens]	94233069 (3107, 3108) Novel Protein sim. GBank gij3043692[dbj BAA25510] - (AB011156) KIAA0584 protein [Homo sapiens]
94233065 (3095, 3096)		32201907 (3099, 3100) Novel Protein sim. Gi gil544463lsplP35350 GUSTATORY RECEI		87617114 (3103, 3104)	94725512 (3105, 3106)	94233069 (3107, 3108) (
1548						1554

1555 8	87332970 (3109, 3110)	87332970 (3109, 3110) Novel Protein sim. GBank gil2257495 dbj BAA21392 - (ABANA54) aidt f (Schirosaccharomycas nombal	Contains protein domain (PF00400) - UNCLASSIFIED WD domain G-bela reneat		264259, 29331826, 35696052, 264508, 264905, 264906, 264907, 264908, 52644045,
					264909, 264910, 60432229, 60433356,
					23612036, 204736, 264739, 33637064, 265011, 87168559, 264601, 265018, 265019,
					264763, 264764, 264288, 264766, 264768.
					21906765, 35695917, 265022, 264691,
				•	33657023, 35696423, 35695855, 264635,
					264555, 264636, 264638, 264639, 18108385,
					56526486
1556	91229268 (3111, 3112)		i	UNCLASSIFIED	83373044, 264758, 265022, 264600,
					35696052, 264630, 35696423, 265018,
					264632, 264682, 29331822, 265020, 265011
1557	87640609 (3113, 3114)	87640609 (3113, 3114) Novel Protein sim. GBank gi 3329611 (AF078783) -	Contains protein domain (PF00097) - UNCLASSIFIED	UNCLASSIFIED	60432289, 264509, 264906, 264907, 264908,
		contains similarity to C3HC4-type zinc fingers (Pfam; zf-	Zinc finger, C3HC4 type (RING		264909, 264910, 264758, 55811386, 264761,
		C3HC4.hmm, score; 34.08); most similar to drosophila	finger)		264762, 264766, 264769, 264690, 263978,
	-	goliath protein (SW: Q06003) [Caenorhabditis elegans]			264634, 264635, 264639, 264564, 264486
1558	94840376 (3115, 3116)	94840376 (3115, 3116) Novel Protein sim. GBank			22278994, 22278996, 22278997, 22278998,
		gi[5360105 gb]AAD42871.1 AF15510 - (AF155105) putative			22278999, 60432049, 264259, 29331824,
		zinc finger protein NY-REN-34 antigen [Homo sapiens]			29331825, 29331826, 29331827, 264908.
					264909, 60433356, 21906754, 265017,
					265018, 264448, 21906767, 265021, 265022,
					33657023, 33657109, 18108370, 55811576,
					83373044, 87168518, 22279000, 22279002
1559	88224865 (3117, 3118)	88224865 (3117, 3118) Novet Protein sim. GBank	Contains protein domain (PF00560) - glycoprotein	glycoprotein	264259, 29331822, 60432289, 35696052,
		gij112908jspjP02750jA2GL_HUMAN - LEUCINE-RICH	Leucine Rich Repeat		264107, 264110, 21906754, 33109954,
	_	ALPHA-2-GLYCOPROTEIN (LRG)			87168559, 264760, 264763, 21906764,
					21906765, 21906769, 265021, 264690,
	_				35695855
1560	84580675 (3119, 3120) Novel Protein sim. GF	Novel Protein sim. GBank gij3880146 emb[CAA92704 -		UNCLASSIFIED	264908, 264603, 264638
		(268319) Similarity to Human hnRNP F protein (PIR Acc.			
		No. S43484); cDNA EST EMBL:D34218 comes from this			
		gene; cDNA EST EMBL: D37248 comes from this gene;			
		cDNA EST EMBL:D71817 comes from this gene; cDNA			
-+	COPPO POPO CAROCOCO	EST EMBL:D74531 comes fro			284510 264504
$\overline{}$	86609139 (3121, 3122)				200001
				UNCLASSIFIED	70330/
				UNCLASSIFIED	264910, 264764, 264766
1564	87766371 (3127, 3128)	Novel Protein sim. GBank	Contains protein domain (PF00441) - dehydrogenase	dehydrogenase	18108394, 35696286, 264259, 29331822,
		gil1168287[sp]P45953JACDV_RAT - ACYL-COA	Acyl-CoA dehydrogenase		60432289, 35696052, 29331828, 264508,
		DEHYDROGENASE, VERY-LONG-CHAIN SPECIFIC			66712502, 264908, 56182435, 265007,
		PRECURSOR (VLCAD)			264910, 60170831, 21906754, 265011,
					265017, 265018, 265019, 18108351, 264448,
					264288, 264686, 21906765, 21906768,
					265021, 60170615, 264692, 35696423,
					35695855, 264557, 56182323, 60432113,
					22279002, 264462

156	67783381 (3129, 3130)	1565 87783381 (3129, 3130) Novel Protein sim. GBank gi 129726 sp P05307 PD _BOVIN - PROTEIN DISULFIDE ISOMERASE PRECURSOR (PDI)		isomerase	264488, 264689, 18108398, 55811957, 264534, 264259, 264508, 264905, 264509,
		(FROLTL 4-HTDROATUSE BEIN SUBUNI) (CELLULAR THYROID HORMONE BINDING PROTEIN) (P55)			264908, 18108372, 264510, 264511, 264512, 265008, 264630, 265009, 264910, 264635,
					264636, 264591, 264555, 264592, 264637.
					264593, 264594, 264595, 264596, 265011,
					264603, 22279002, 18108351, 264762,
1568	87424740 /2121 2122	Marie Destrict of the Court all population and the control			264565, 264567
<u> </u>	(2010, 1010) 84742410	0.144749 (5.151, 5.152) NOVELFFORM SIM. GBBIR gij368U445 emb CAAZU329		į	22278996, 22278999, 264259, 29331822,
		(ALUS IZDO) VM IUDR. I [Caendrhabditis elegans]			29331824, 60432289, 29331827, 66712502,
_					264908, 265008, 18108351, 52644229,
					21906765, 21906767, 21906768, 21906769,
1587		None Details are Object			33657109, 264555, 264639, 264482
<u> </u>		otagacco (5155, 5154) Novel Protein Sim. Chank		UNCLASSIFIED	56182575, 21906769, 264692
		gi(4929699)gb/AAD34110.1(AF15187 - (AF151873) CGI-115			
3		protein [Homo sapiens]			
1568	87648761 (3135, 3136)	Novel Protein sim. GBank	Contains protein domain (PF00096) - Iranscriptfactor		29331827, 29331830, 264511, 265009,
		gil4827063 ref NP_005072.1 pZNF1 - zinc finger protein 142	Zinc finger, C2H2 type		264758, 21906767, 21906768, 264691,
		(clone pHZ-49)			264693, 22279000, 22279002
1569	90936668 (3137, 3138)	90936668 (3137, 3138) Novel Protein sim. GBank gi 5689451 db BAA83009.1 -	Contains protein domain (PF00443) - ubiquitin	ubiquitin	65274572, 29331822, 29331824, 29331828,
		(AB028980) KIAA1057 protein [Homo sapiens]	Ubiquitin carboxyl-terminal hydrolase		264905, 56182435, 265007, 265019, 264764,
			family 2		21906765, 21906769, 55811957, 60170615,
					52644150, 264692, 33657023, 33657109,
					18108377, 264563, 264567
1570		86943981 (3139, 3140) Novel Protein sim. GBank gij1255430 (U53155) - No Idefinition line found (Caencharditis aborate)		UNCLASSIFIED	264595, 264682, 265021
1571			Canada distanta	to the call in	Logical Fountain Sections Sections
<u> </u>		CONTRACTOR OF A CONTRACTOR OF	Contains protein domain (Produst) - Ideulin		ZZZ/8996, 35696286, ZZZ/8997, Z64091,
_		gilebor / 3 ifretine _ but bot : I pri bec - tubulin, gamma	l ubulin/FtsZ tamily		264259, 29331824, 29331825, 29331827,
		polypeptide			35696052, 264508, 264905, 56182435,
					264510, 265007, 264758, 265011, 18108351,
					264448, 264288, 264369, 21906765,
					21906767, 21906768, 21906769, 35695917,
					265020, 265021, 33657023, 264693,
					18108370, 18108377, 35696423, 35695855,
	_				264634, 264555, 264558, 18108384
15/2	81201664 (3143, 3144)			UNCLASSIFIED	35696052, 264905, 264906, 264908, 264910,
					264758, 264766, 35695917, 264637
13/3	80207066 (3145, 3146)			UNCLASSIFIED	263972

gij2881810jemb CAA94856 - dd calcium binding protein; cDNA s from this gene [Caenorhabditis] from this gene [Caenorhabditis] gij2773195 (AF039711) - comitrella patens glyceraldehyde se (GB:X72381) [Caenorhabditis] SP_MOUSE - WISKOTT - ROTEIN HOMOLOG (WASP) gij2978255[dbjjBAA25190] - inger protein-2 [Mus musculus] gij4324682[gbjAAD16986] - lung protein 1 [Rattus norvegicus] gij20888338 (AF003386) - senorhabditis elegans]	Contains protein domain (PF00173) - Cytochrome 18108394, 264887, 18108397, 18108398, Heme-binding domain in cytochrome 22278996, 22278996, 22278999, 264259, 264259, 26331827, 29146499, 26431827, 29146499, 2644107, 264907, 264909, 52644045, 264511, 265009, 264909, 2644045, 264511, 265009, 264909, 2644045, 264511, 265019, 18108351, 264682, 264763, 264763, 264686, 21906765, 264086, 21906765, 21906765, 21906766, 21906766, 21906769, 264691	Contains protein domain (PF00036) - phosphatase 56994075, 264259, 29331822, 29331824, 29331825, 60432289, 29331822, 29331824, 29331825, 60432289, 29331828, 264909, 265006, 265008, 265009, 265010, 87168559, 26511150, 264448, 18108354, 264369, 265020, 265021, 60170615, 264691, 33657023, 33657109, 60431528, 65274791, 35695855, 18108385, 60432113, 22279002, 26482	UNCLASSIFIED 264569, 264092, 264094, 264095, 264259, 264509, 264509, 264509, 264509, 264909, 264509, 264509, 264910, 264909, 264510, 264511, 265006, 264910, 21906754, 265001, 265011, 87168559, 264761, 264761, 264762, 264691, 264693, 35695855, 264632, 264634, 264693, 83373044, 264468	UNCLASSIFIED 264511, 265011, 264681, 264369, 264686, 264689, 264629, 264555, 264558, 264559 264559, 264559, 264559 264693	ĒŌ	Contains protein domain (PF00188) - glycaprotein 22278999, 35696052, 29331830, 52644045, SCP-like extracellular protein 265022, 26438, 87168474, 265018, 264448, 265022, 26438, 56526486, 22279000 22278995, 29331822, 29331824, 29331826, 263017, 264288, 21906764, 55811957, 266017, 264288, 21906764, 55811957, 266017, 264288, 21906764, 55811957, 266017, 264288, 21906764, 55811957, 264018, 26411957, 264018, 26411957, 264018, 26411957, 264018, 26411957, 264018, 26411957, 264018, 26411957, 264018, 26411957, 264018, 26411957, 264018, 26411957, 264018, 26411957, 264018,
95340019 (3147, 3148) 95340019 (3149, 3150) 95314019 (3151, 3152) 87613800 (3153, 3154) 87123138 (3155, 3156) 88085141 (3157, 3158) 87255702 (3159, 3160)		56l - : cDNA abditis	?773195 (AF039711) - nirella patens glyceraldehyde (GB:X72381) [Caenorhabditis	_MOUSE - WISKOTT- TEIN HOMOLOG (WASP)		rvegicus]
	4216142 (3147, 3148) Novel Protein sim. GBank gl 4758334 ref NP_004256.1 g desaturase	15340019 (3149, 3150) Novel Protein sim. GBank gil3 (Z70783) similar to EF-hand o EST EMBL.C08700 comes fro elegans]	15314019 (3151, 3152) Novel Protein sim. GBank gilg contains similarity to Physcorr 3-phosphate dehydrogenase elegans]	7613800 (3153, 3154) Novel Protein sim. GBank gilz499130 sp P70315 WASP ALDRICH SYNDROME PROT 7123138 (3155, 3156)	8085141 (3157, 3158) Novel Protein sim. GBank gilz (AB007407) myeloid zinc fing	17255702 (3159, 3160) Novel Protein sim. GBank gild (AF109674) late gestation lun 15087431 (3161, 3162) Novel Protein sim. GBank gild F59E12.4 gene product [Caer

264259, 60432289, 29331627, 264509, 264905, 264906, 264907, 264909, 264910, 264762, 264289, 264768, 264769, 264632, 264555, 264639, 56526486, 22279000	60170831, 33657402, 264682, 21906766, 35695855, 264563	60424179, 52646842, 65274572, 56182575, 22278996, 22278996, 22278998, 22278999, 264259, 29331822, 56182181, 60424269, 60432289, 29331822, 56182181, 60424269, 60432289, 29331822, 29331830, 35696052, 29346498, 66712502, 29331830, 2564046, 56182435, 264510, 264512, 265008, 6043336, 33557402, 60433438, 55812036, 21906754, 55811386, 52644296, 87168474, 87168559, 265018, 266019, 264448, 264369, 265021, 265022, 52644150, 33567021, 265022, 52644150, 33557023, 33557109, 18108374, 52610764, 55810764, 5696423, 65274791, 35655865, 56182273, 83737944, 18108387, 87188518, 60432113, 22279002	35696286, 22278998, 264259, 29331822, 29331824, 29331824, 29431824, 265006, 265007, 265008, 60433356, 33109954, 87168474, 265011, 265017, 264017, 264017, 264017, 264017, 264017, 264056, 18108369, 18108359, 21906765, 18108384, 18108388, 264482, 263972, 18108383, 18108388, 264482, 264564	265017, 265018, 264689, 33657023, 263978, 264636, 264563	264907, 264908, 264511, 264910, 264591, 264594, 264629, 264631, 264563, 264483, 264567	264259, 29331828, 264905, 265006, 264758, 21906754, 264761, 264762, 21906765, 21906769, 60170615, 52644150, 33657109, 35695855, 56182323, 18108385	65274572, 264490, 29331822, 66714117, 29331827, 29331828, 56182435, 265008, 60170831, 264595, 264758, 264596, 265011, 264686, 21906766, 21906768, 55811957, 27486265, 264639, 18108385, 56526486, 60432113
phosphalase	UNCLASSIFIED	dehydrogenase	phosphatase		UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED
		Contains protein domain (PF00106) - dehydrogenase short chain dehydrogenase	Contains protein domain (PF00023) - phosphatase Ank repeat			Contains protein domain (PF00468) · Ribosomal protein L34	
1582 95358052 (3163, 3164) Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]	87622715 (3165, 3166) Novel Protein sim. GBank gi[5578956 emb CAB51351.1 - (AL050306) dJ475B7.2 (novel protein) [Homo sapiens]	95337722 (3167, 3168) Novel Protein sim. Gank gl5531815[gb]AAD4482.1 - (AF078850) steroid dehydrogenase homolog [Homo sapiens]	87626117 (3169, 3170) Novel Protein sim. GBank gil4240132 dbijBAA74846.1 - (AB020630) KIAA0823 protein [Homo sapiens]	88067081 (3171, 3172) Novel Protein sim. GBank gij3786494 (AF098993) - No definition line found [Caenorhabditis elegans]	87617126 (3173, 3174) Novel Protein sim. GBank gi[3253159 (AF005355) - translation initiation factor elF2C [Oryctolagus cuniculus]	Novel Protein sim. GBank gi 1077573 pir S52680 - probable Contains protein domain (PF00468) - UNCLASSIFIED ribosomal protein L34 Ribosomal protein L34 Saccharomyces cerevisiae)	90980653 (3177, 3178) Novel Protein sim. GBank gij2137756 pir 48746 - semaphorin C - mouse (fragment)
95358052 (3163, 3164)	_		87626117 (3169, 3170)			87802536 (3175, 3176) Novel Protein sim. Giribosomal protein L34 (Saccharomyces cere	90980653 (3177, 3178)
1582	1583	1584	1585	1586	1587	1588	1589

1980	1590 95319825 (3179, 3180)			UNCLASSIFIED	293489, 22278996, 264259, 29331824, 29331824, 29331825, 29331827, 265006, 6043356, 21906754, 265017, 265018, 26619, 264448, 264765, 264288, 5264229, 21906765, 21906767, 21906768, 21906769, 26486, 60432113, 22279000, 22279002
$\overline{}$	86877160 (3181, 3182)				264564
	87882533 (3183, 3184) Novel Protein sim. Gl gil4557749 ref NP_0C transactivator	Novel Protein sim. GBank gil4557749 ref NP_000237.1 pMHC2 - MHC class II transactivator		мнс	264259, 264905, 29331830, 264595, 265017, 264448, 264288, 264690, 264629, 87168518
	94991661 (3185, 3186)			UNCLASSIFIED	65274572, 60432049, 264509, 60433356, 21906754, 21906767, 21906768, 18108370, 34606437, 2272010, 264468, 264657
1594	87773752 (3187, 3188)	87773752 (3187, 3188) Novel Protein sim. GBank gi[3877072]emb[CAAB7060] -		UNCLASSIFIED	264488, 29331827, 264905, 264906, 264907
-		(Z46937) similarity with ribosomal protein L21			264908, 264909, 264910, 264592, 264593,
					264757, 264602, 264604, 264760, 264681, 3 264288, 264766, 264768, 29148629
					35695917, 264692, 264628, 264629, 264630,
					264632, 264634, 264635, 264636, 264639, 264563, 264564, 264566
1595 7	79919425 (3189, 3190)	79919425 (3189, 3190) Novel Protein sim. GBank gij3152703 (AF065389) -	Contains protein domain (PF00335) - UNCLASSIFIED		29331826, 264908, 55811957
		tetraspan NET-4 [Homo sapiens]	4 transmembrane segments integral		
_	70033028 /3101 3102		membrane proteins	ı	
	36074087 (3131, 3132)			SSIFIED	29146498, 264758, 263967
	009/165/ (3195, 3194)	dos/100/ (3193, 3194) Novel Protein sim. GBank	Contains protein domain (PF00067) - cyto450		264092, 29331824, 264508, 264682, 264369,
		gijozo 714 gblaAD41244. IJAF09448 - (AF094480) cholesterol 24-hydroxylase [Homo sapiens]	Cytochrome P450		264686, 264630, 264563
1598	87862939 (3195, 3196)	87862939 (3195, 3196)			264259, 264634
1599	87649829 (3197, 3198)	Novel Protein sim. GBank	Contains protein domain (PF01581) - UNCLASSIFIED		52645080, 29331824, 29331826, 264511,
		gil4506797[ref[NP_000324.1]pSCA7 - spinocerebellar ataxia FMRFamide related peptide family 7 (aliscontropycholler strate, with posterior)	FMRFamide related peptide family		265009, 265011, 264605, 264448, 264764,
		(University and all opin) with retinal degeneration)			265020, 264692, 264693, 18108370, 264635, 18108385
909	80056002 (3199, 3200)				29331826, 264603, 264691, 264563
-	15023246 (3201, 3202)	15023246 (3201, 3202)		ASSIFIED	264635
	56926987 (3203, 3204)	Novel Protein sim. GBank oii5305704lob AAD41780 11AE12853 - (AE128535)	Contains protein domain (PF00018) - struct		29146499, 264112, 264762, 18108351,
		cytoplasmic phosphoprotein PACSIN2 [Mus musculus]	one demail		29148627, 263974
1603	80502072 (3205, 3206)	80502072 (3205, 3206) Novel Protein sim. GBank gil 283920 pir S27939 - tensin -		collagen	264490, 29331824, 264907, 264909, 264511,
		cnicken			265008, 264592, 265010, 265011, 264762,
					264764, 264369, 264288, 264687, 264769,
				•	ZD4833, Z04628, Z64634, Z64636, Z64555,
				<u> </u>	204336, 204630, 204337, 204336, 264339, 18108385
9	80221813 (3207, 3208)	80221813 (3207, 3208) Novel Protein sim. GBank gil4768831 gb AAD29633.1 AF11682 - (AF116827) unknown Homo sanians		ATPase_associated 263977	263977
		cupide control			

1006	160E 104224430 (2200 2240)			struct	264905, 264509, 264906, 264907, 264908,	_
3	31221123 (3203, 3210)			·	264909, 264604, 264766, 264768, 264692,	
					264693, 33657109, 264629, 35695855,	_
					264635, 264636, 264637	
1606	94312703 (3211, 3212)	94312703 (3211, 3212) Novel Protein sim. GBank	Contains protein domain (PF00047) - struct		22278996, 22278999, 264259, 33657402,	
		gi 4505313 ref NP_003794.1 pMYOM - UNKNOWN	Immunoglobulin domain		255017, 18108351, 264446, 21900707, 21906769, 52644150, 264691, 87168518	
1607	10871805 (3213, 3214)	10871805 (3213, 3214) Novel Protein sim. GBank		transcriptfactor	264689	
		gij5174473 ref NP_005888.1 pIPPI - intracisternal A particle- promoted polypeptide				
1608	_	80428900 (3215, 3216) Novel Protein sim. GBank gij2224629jdbj BAA20802 -		UNCLASSIFIED	264094, 264906, 264907, 264909, 264910, 264591, 264603, 264768, 264693, 264634	
					264635, 264637, 264639	_
1609		94311572 (3217, 3218) Novel Protein sim. GBank gil4884073 emb CAB43213.1 -			52644507, 52645156, 52646365, 52646842,	
		(AL049934) hypothetical protein [Homo sapiens]			56182575, 22278994, 56994075, 35696286,	
					22278997, 22278998, 22278999, 264259,	
					52645080, 29147620, 29331826, 35696052,	
					33656970, 264508, 264509, 264907,	_
					52644045, 56182435, 264510, 264511,	
					264512, 33657402, 21906754, 52646317,	_
					33109954, 52644296, 87168474, 265017,	
					265018, 265019, 18108351, 264448, 264288,	
					264769, 52644229, 21906765, 21906766,	
					21906767, 21906768, 21906769, 55811957.	
					35695917, 265020, 265021, 265022,	
_					52644150, 33657023, 33657109, 52645129,	
_					27486261, 27486262, 35695763, 264628,	_
					18108370, 18108376, 35696423, 264638,	
					52644332, 18108387, 87168518, 22279000,	
				011004	204303, 204400 201503 201707 0012030 205019 205030	7
1610	85468200 (3219, 3220) Novel Protein sim. G	Novel Protein sim. GBank gi[283920 pir] 527939 - tensin -		UNCLASSIFIED	204593, 204737, 33012030, 203010, 203020, 204691, 204691, 264694, 264693, 2646931, 2646934,	
		CHOCKET			264635, 264555, 22279000, 264564	_
1611	7	94122843 (3221, 3222) Novel Protein sim. GBank gi 107284 pir A35415 -	Contains protein domain (PF00008) - peroxidase	peroxidase	35696286, 21906765, 264691, 35696423	
		peroxidase (EC 1.11.1.7), thyroid (Grave's disease) - human EGF-like domain	n EGF-like domain			
		(fragment)			20000	Т
1612	85746031 (3223, 3224)	ein sim. C		UNCLASSIFIED	264488, 264509, 18108370, 18108387.	
					26448b	-
		protein (SW.TCR2_BACSU); cDNA EST EMBL:C09951				
		comes from this gene; cDNA EST EMBL:C08265 comes				
		from this gene [Caenorhabditis elegans]		1	,	Т
1613	1613 82247354 (3225, 3226)			UNCLASSIFIED	264759	٦

1614	191228634 (3227 3228)	1614 191228634 (3227 3228) Novel Protein sim GBank	Contains protein domain (PEO1605) - [LINC] ASSIETED	INCI ASSIEIED	92278004 22278006 22278007 2278008
		gi 4680673 gb AAD27726.1 AF13295 - (AF132951) CGI-17	eRF1-like proteins		22278999, 264259, 29331822, 264908,
		protein [Homo sapiens]			264512, 265009, 265011, 265017, 265018,
					265019, 18108351, 264683, 264288, 264766,
					21906767, 21906768, 21906769, 35695917,
					265021, 265022, 35696423, 35695855,
4646		Manual Destate and OBack alternations of the Control of the Contro			60170394, 56182323, 83373044, 264566
2		oo iz isus (szzs, szsu) Novel Protein sim. Gbank gijsbass485 dbjjBAA83UZ6.1 -	Contains protein domain (PF 00023) - homeobox	нотеорох	22278996, 35696286, 22278997, 29331822,
		(Abuzoss/) NIAA IU/4 protein [Homo sapiens]	Ank repeat		35696052, 29331828, 264508, 264908,
					264909, 56182435, 264511, 265017, 265019,
					264766, 264767, 264768, 265020, 264691,
					264628, 264632, 264635, 264555, 264556,
1616		94311819 (3231 3232) Novel Protein sim (38ank gil3876260lemblC&B046961		INC. ACCIEIED	30102323, 204338, 2227930Z
2		(278418) cDNA EST EMBL:D71020 comes from this gene:		ONCLASSIFIED	204408, 32044307, 32043 (30, 32040303, 32040803, 32040803, 32040842, 22278994, 22278995, 35696286,
		cDNA EST EMBL:D73593 comes from this gene; cDNA			22278996, 22278997, 22278999, 52645080,
		EST EMBL: C07649 comes from this gene; cDNA EST			29331822, 29331824, 29331825, 29331827,
		EMBL:C09081 comes from this gene; cDNA EST yk399f2.3			29331828, 35696052, 33656970, 264905,
		comes from this gene; cDNA			264909, 264594, 52646317, 21906754,
					33657084, 52644296, 87168474, 87168559,
					265017, 265018, 265019, 264681, 264448,
					264684, 52644229, 21906764, 264689,
					21906765, 21906766, 21906769, 35695917,
					265020, 265021, 52644150, 33657023,
					52645129, 33657109, 33657182, 27486261,
					27486262, 33657349, 27486265, 35695763,
_					18108376, 35696423, 35695855, 264557,
	_				52644332, 264558, 18108385, 87168518
1617	88090742 (3233, 3234) Novel Protein sim. (Novel Protein sim. GBank	Contains protein domain (PF01529) - peptidase	peptidase	35696052, 264905, 264509, 264907, 264908,
		gil466053[sp]P34679[YO41_CAEEL - HYPOTHETICAL 68.7]DHHC zinc finger domain	7 DHHC zinc finger domain		264510, 264511, 264764, 264766, 264768,
		KD PROTEIN ZK757.1 IN CHROMOSOME III			264689, 264693, 18108374, 264635, 264636,
	_				264638
9191		862/2860 (3235, 3235) Novel Protein sim. GBank gil4240231 dbj BAA74894.1 -	•	struct	35696286, 22278999, 264092, 29331824,
	_	(ABUZU678) KIAAU871 protein [Homo sapiens]			29331825, 35696052, 33657084, 21906765, 27486264
1619	95354580 (3237, 3238) Novel Protein sim.	Novel Protein sim. GBank	Contains protein domain (PF00010) - transcriptfactor	transcriptfactor	52646842, 65274572, 22278999, 264259.
		gi 5031763 ref NP_005515.1 pHRY - hairy (Drosophila)-	Helix-loop-helix DNA-binding domain	•	29331822, 29331824, 29331825, 29331826,
			•		29331827, 29331828, 35696052, 56182435,
					265007, 265008, 264910, 60170831,
					60432229, 60433356, 60433438, 265019,
					264448, 264288, 264686, 21906768, 265021,
					60170615, 33657023, 65274620, 33657109,
					18108374, 18108376, 35696423, 35695855,
000	10100 0000 33011000				56182323, 56526486
0291	87344655 (3239, 3240)	1020 87344655 (3239, 3240) Novel Protein sim. GBank gein314715p P45843 SCRT_DROME - SCARLET		UNCLASSIFIED	264684
		PROJEIN			

1621	1621 87076708 (3241, 3242)					
1622	94741739 (3243, 3244)	94741739 (3243, 3244) Novel Profess sim Grank nitracos Anna Anna Continued Continued to Continue		UNCLASSIFIED	264910	
		calmodulin-binding protein [Gallus gallus]	Contains protein domain (PF00622) - SPRY domain	UNCLASSIFIED	18108392, 65274572, 18108398, 22278996, 22278997, 22278999, 29146498, 29146499, 264905, 264908, 264909, 264828, 52644045, 264592, 60433356, 21906764, 284602, 265017, 264389, 21906768, 55811957, 265021, 60170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 2272000, 224565	
1623		87779106 (3245, 3246) Novet Protein sim. GBank gij731086 sp P40389 UV22_SCHPO - UV-INDUCED PROTEIN UVI22		ribosomalprot	18108398, 264259, 264909, 56182435, 18108398, 264269, 264909, 56182435, 26464, 2746648, 21906768, 35695917,	
1624		87338178 (3247, 3248) Novel Protein sim. GBank gij3875666[emb]CAB05478] - (283104) cDNA EST EMBL: T00015 comes from this gene; cDNA EST EMBL: D33665 comes from this gene; cDNA EST EMBL: D36540 comes from this gene; cDNA EST 9/k240f8.3 comes from this gene; cDNA EST 9/k240f8.3 comes from this gene; cDNA EST 9/k387c8.3 comes from this gene; cDNA EST 9/k387c8.3		UNCLASSIFIED	ZE4091, 8/105916, Z64503 66714117, 29331825, 264909, 265008, 264758	
1626		94734369 (3251, 3252) Novel Protein sim. GBank gli4589622[db]BAA76833.1 - gli5679070[gb]AAD46844.1 AF16090 - (AF160904) BCDNA.HL05936 [Drosophila melanogaster]		kinase	264489, 22278994, 22278996, 35696296, 22278997, 22278998, 22278999, 264092, 2642897, 22278998, 22278999, 264092, 2642897, 22278998, 22278999, 264092, 264289, 29331825, 264308, 264102, 264106, 264008, 264208, 264318, 264017, 265011, 264011, 265011, 264011, 265011, 264011, 265011, 264011, 265011, 264011, 265011, 264011, 265011, 264011, 265011, 264011, 26274620, 18108368, 263972, 18108376, 3696423, 26278994, 56994075, 22278997, 22278998, 5694075, 22278997, 22278999, 6043209, 265009, 265009, 265009, 265009, 265011, 87168559, 264604, 265011, 87168559, 264604, 265011, 266021, 266021, 266011, 266011, 266021, 266021, 265011, 27486262, 27486261, 2654332, 263972, 2524433, 263972, 25244332, 263972, 27486262, 27486261, 264332, 263972, 27486262, 27486262, 27486262, 27486261, 264332, 263972, 2524332, 263972, 27486262, 27486262, 27486261, 266021, 266001, 266001, 266001, 266001, 266001, 266001, 266001, 266001, 2660	
1627				UNCLASSIFIED	264288	
1628	85708459 (3255, 3256) Novel Protein sim. GB hypothetical protein [A	Novel Protein sim. GBank gij3668087 (AC004667) - hypothetical protein [Arabidoosis thaliana]		eph	264288, 264686, 264767, 22279002	
		forman rodona. J				

264555	29331822, 29331827, 265010, 264693, 264634, 22279002	55811957, 264259, 33657023, 264693, 29331824, 29331824, 29331827, 29331828, 264906, 55811576, 264910, 264639, 264636, 26482323, 264559, 264558, 264588, 264584, 264588	16103503, 784303, 284704, 264708 264687, 264769, 264691, 264692, 29146499, 264509, 264905, 264907, 264511, 264512, 264487, 264581, 264573, 264682, 264682	264488, 264259, 264907, 264908, 264909, 264628, 264629, 264803, 264909,	265007, 284637, 22279002	65274572, 22278998, 35696052, 52644045, 264511, 265008, 265009, 265010, 265011, 265019, 264448, 264369, 21906765,	22278995 264594 264763 265020	264558	21906765, 21906767, 22278996, 35696286, 22278999, 264259, 264692, 264693, 29331824, 33657109, 264508, 264906, 18108370, 264629, 265007, 33657402, 21906754, 264602, 264604, 264764, 264683, 264568, 264508	264488, 18108394, 65274572, 56182575, 264488, 18108394, 65274572, 56182575, 35696286, 29331824, 29331826, 29331827, 35696052, 264908, 56182435, 264712, 265006, 265008, 264757, 264758, 55811386, 264603, 264766, 18108351, 264764, 264788, 264768, 29108767, 55811957, 264691, 33657023, 65274620, 18108370, 55810764, 55811957, 264518, 264538, 264539, 2645318,
struct	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED		struct	UNCLASSIFIED	UNCLASSIFIED	kinase		UNCLASSIFIED
					Contains protein domain (PF00435) - struct Spectrin repeat			Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain		
Novel Protein sim. GBank gil4240175 dbj BAA74866.1 - (AB020650) KIAA0843 protein [Homo sapiens]		Novel Protein sim. GBank gji1913892 emb CAA69995 - (Y08740) tom-1A protein [Gallus gallus]	Novel Protein sim. GBank gi[2558501 dbj BAA22896 - (D63850) hepatoma-derived growth factor [Mus musculus]		Novel Protein sim. GBank gil4887229[gb AAD32244.1 AF15075 - (AF150755) microtubule-actin crosslinking factor [Mus musculus]			Novel Protein sim. GBank gil4557511frefINP_001339.1 pDAPK - death-associated protein kinase 3	87101854 (3275, 3276) Novel Protein sim. GBank gij3420051 (AC004680) - unknown protein [Arabidopsis thaliana]	94322194 (3277, 3278) Novel Protein sim. GBank gil5420389lemb[CAB46680.1] - (AJ243460) proteophosphoglycan [Leishmania major]
1629 84993841 (3257, 3258) Novel Protein sim. (AB020650) KIAA0	87779027 (3259, 3260)	87758454 (3261, 3262) Novel Protein sim. (Y08740) tom-1A p		87773683 (3265, 3266)		1035 54232000 (3608, 3270)	80413227 (3271, 3272)	80070435 (3273, 3274) Novel Protein sim. gil4557511freflNP_ protein kinase 3	87101854 (3275, 3276)	94322194 (3277, 3278)
1629	1630	1631		1633	1634	650			1638	1639

1640	94143185 (3279, 3280)	1640 94143185 (3279, 3280) Novel Protein sim. CBank gi[2842469lemb CAA16847.1 - (AL021747) hypothetical protein [Schizosaccharomyces pombe]			56182575, 56994075, 35696286, 60432049, 60432289, 29331827, 35696052, 52644045, 56182435, 264510, 265006, 265007, 265008, 265010, 265011, 265017, 265018, 2644229, 21906765, 21906766, 21906766, 21906766, 21906768, 35695917, 265021, 60170615, 52644150, 33657023, 33657091, 35695855, 264632, 264655, 56182323, 22279000
1641	87625160 (3281, 3282)		1	UNCLASSIFIED	29146499, 265006, 265007, 265008, 265009, 55812038, 265010, 265011, 264555, 264556, 264558, 2645888, 2645888, 2645888, 264588, 264588, 264588, 264588, 264588, 2645888, 2645888, 264588, 26458
1642	94312557 (3283, 3284)	94312557 (3283, 3284) Novel Protein sim. GBank gi 1575333 (U60416) - myr 6 myosin heavy chain [Rattus norvegicus]	Contains protein domain (PF01843) - struct DIL domain		22278999, 29147620, 29331826, 29331828, 33656970, 55812038, 265010, 265018, 265019, 18108351, 264689, 265020, 265022, 264690, 33657023, 65274620, 35695763, 52644332, 18108381, 60170394, 56182323, 18108388, 87168518, 22279002, 264564
1643	94131766 (3285, 3286)				29331625, 29331627, 29331628, 21906754, 265019, 264288, 264693, 33657349, 18108370, 18108376, 264555, 83373044, 22279002, 264462
1644	88095125 (3287, 3288)			UNCLASSIFIED	264905, 264907, 264908, 264910, 265009, 264757, 264758, 264761, 264762, 264763, 264766, 264768, 264628, 264630, 264631, 264563, 264565, 264566, 264566, 264566, 264566, 264567
1645	95013858 (3289, 3290)			UNCLASSIFIED	264685, 264693
1646	95362691 (3291, 3292) Novel Protein sim. G	Novel Protein sim. GBank gil 1076802 pir S49915 - extensin like protein - maize		UNCLASSIFIED	22278994, 56994075, 35696286, 264259, 29331824, 29331825, 29331826, 60432289, 264508, 60433356, 60433438, 87168559, 265018, 264687, 35695917, 264692, 33657023, 33657182, 27486261, 27486265, 33657349, 60432113, 264563, 264564
1647		94278428 (3293, 3294) Novel Protein sim. GBank gi 5002573 emb CAB44338.1 - (Y17466) alpha-N-acetylgalactosamine alpha-2.6- slalyltransferase [Fugu rubripes]		UNCLASSIFIED	29331822, 264906, 264908, 264369, 21906768, 60170615, 264639, 22279000
1648	87642098 (3295, 3296)		Contains protein domain (PF00096) - UNCLASSIFIED Zinc finger, C2H2 type	UNCLASSIFIED	265009, 264686, 55811957, 35955917, 55810764, 264556, 56182323, 264558, 18108385

264488, 22278995, 35696286, 22278996, 22278997, 22278997, 22278998, 60432049, 264259, 29331822, 29331822, 29331822, 29331826, 29331822, 29331827, 29331826, 29331827, 29331827, 264908, 264908, 264908, 264908, 264908, 264908, 264908, 264908, 264908, 264691, 264601, 264602, 265017, 265018, 265019, 264682, 264002, 265017, 265018, 265019, 264682, 264687, 264683, 264688, 264686, 264682, 264682, 264689, 21906766, 21906767, 21906768, 55811957, 35695817, 265021, 265022, 264639, 20281149, 18108370, 264632, 264632, 264634, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 18108376, 264044, 60432113, 22279000, 22279002, 264482, 264565, 264487, 26279000, 22279002, 264482, 264565, 264487, 264565, 264566, 264487, 264566, 264487, 264565, 264566, 264487, 264565, 264566, 264487, 264566, 264487, 264565, 264566, 264487, 264566, 264487, 264567, 264687, 2646	265011, 264602, 21906767, 18108374, 18108377, 18108385	264488, 52645156, 18108397, 35696286, 22278998, 224259, 29331822, 29331824, 29331825, 29331826, 29331827, 264910, 60432229, 21906754, 265010, 265011, 265017, 265019, 264288, 264688, 21906765, 21906766, 21906765, 21906766, 21906765, 21906766, 21906765, 264037, 264693, 264657, 264557, 264559, 83373044, 56526486, 22279000, 22279002, 264564	22278997, 29146498, 56182435, 21906754, 264369, 21906765, 21906768, 2190676, 2155020, 52644150, 33657109, 22279000, 22279002	18108398, 22278998, 22478998, 264259, 29331822, 29331824, 60432289, 29331826, 29331827, 29331837, 264909, 265006, 265009, 6043229, 60433356, 60433438, 21906754, 265017, 265022, 264692, 18108364, 665274791, 18108384, 60432113, 264567	33657109, 264565
cadherin		synthase	struct	glycoprotein	UNCLASSIFIED
			Contains protein domain (PF00169) - struct PH domain	Contains protein domain (PF00047) - glycoprotein Immunoglobulin domain	
1649 95347628 (3297, 3298) Novel Protein sim. GBank gil854065 emb CAA58337 -	B7418539 (3299, 3300) Novel Protein sim. GBank gi[3647335 emb CAA21059 - (AL031644) possible zinc-finger protein [Schizosaccharomyces pombe]	91639773 (3301, 3302) Novel Protein sim. GBank gi[4884278]emb[CAB43247.1] - (AL050037) hypothetical protein [Homo sapiens]	86598622 (3303, 3304) Novel Protein sim. GBank gil1657837 (U73200) - p116Rip [Mus musculus]	94255993 (3305, 3306) Novel Protein sim. GBank gi 3776054 emb CAA06273 - (AJ004999) Tapasin [Gallus gallus]	
95347628 (3297, 3298) P	87418539 (3299, 3300)				79756471 (3307, 3308)
1649	1650	1651	1652	1653	1654

1656 16660346 (1309), 3310] Novel Protein sim. GBank gi 335271/emp CAx73496] 1656 16660346 (1309), 3310] Novel Protein sim. GBank gi 335271/emp CAx73496] 1652001 13312, 3312] Novel Protein sim. GBank gi 390041
Synthase Contains protein domain (PF00076) - dna_ma_bind RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) Contains protein domain (PF00829) - UNCLASSIFIED Ribosomal prokaryotic L21 protein UNCLASSIFIED UNCLASSIFIED
synthase UNCLASSIFIED - UNCLASSIFIED UNCLASSIFIED
Q Q
52644507, 35696286, 22278998, 22278999, 29331824, 29331825, 29331828, 33656970, 264908, 52644045, 264511, 264910, 52646317, 264288, 52644229, 33657023, 33657109, 52644322, 2844527, 56182323, 65526486, 60432113, 29331825, 264563, 29331827, 265009, 21906766, 21906767, 265020, 265020, 265020, 265020, 265020, 264289, 264289, 264638, 25278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 264289, 264303, 264638, 21906769, 29331822, 29331822, 29331822, 29331822, 29331822, 29331822, 29331822, 29331822, 2644045, 266006, 60433356, 264757, 2644046, 265021, 18108351, 264408, 264408, 265021, 265022, 18108362, 264408, 265021, 265022, 18108362, 2644097, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264617, 264619, 264619, 264690, 264602, 264602, 264609, 264602, 264609, 264602, 264609, 264602, 264602, 264609, 264602, 264602, 264609, 264602, 264602, 264603, 264602, 264603, 264603, 264603, 264603, 264603, 264603, 264603, 264603, 264603, 264639, 264639, 264639, 264603

UNCLASSIFIED 264488, 22278998, 264259, 29331824, 29331826, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331828, 264509, 66712502, 29331830, 264908, 5264045, 265007, 264512, 60433356, 60433438, 26812038, 21906754, 265019, 264448, 269766, 264768, 264769, 21906768, 264769, 21906768, 264769, 264767, 264563, 264567, 26	18108392, 29331822, 29331828, 20281100, 284106, 265006, 265007, 265008, 18108348, 21906766, 18108365, 18108366, 18108374, 81337344, 48108385		UNCLASSIFIED 264488, 263994, 3569628B, 29331824, 35696052, 264508, 264509, 264905, 264905, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 26509, 264910, 60170831, 264591, 264591, 264591, 264604, 264604, 264604, 264760, 264762, 18108351, 264604, 264604, 264763, 264763, 264764, 264764, 264764, 264764, 264764, 264684, 264764, 264764, 264764, 264769, 264684, 264764, 264591, 264692, 33657109, 33657182, 264632, 3696423, 3696423, 3696423, 264634, 264636, 264636, 264634, 264638, 264636, 264639, 264634, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264556, 264636, 264638, 264638, 264563, 264563, 264568, 264486
Contains protein domain (PF00076) - UNCLASSIFIED RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	Ш	Contains protein domain (PF00397) - kinase WW domain	
Novel Protein sim. GBank gil4759100[ref]NP_004759.1[pSFRS - splicing factor, arginine/serine-rich 11	94135172 (3323, 3324) Novel Protein sim. GBank gil1730502 sp P52875 PF27_MOUSE - TRANSMEMBRANE PROTEIN PFT27	94217146 (3325, 3326) Novel Protein sim. GBank gil4884136 emb CAB43275.1	94234076 (3327, 3328) Novel Protein sim. GBank gij3043692 dbj BAA25510 - (AB011156) KIAA0584 protein [Homo sapiens]
94234071 (3321, 3322)	94135172 (3323, 3324)		1664 94234076 (3327, 3328) N

264488, 29331826, 29331828, 264509, 264906, 264907, 264909, 264510, 264910, 264592, 264593, 264595, 264758, 264596, 264760, 264760, 264762, 264766, 264766, 264766, 264766, 264766, 264766, 264766, 264636, 264636, 264636, 264567, 264486	56994075, 22278999, 264259, 29331824, 29331826, 29331826, 29331827, 29146498, 265009, 33109954, 87168559, 265019, 264288, 269688, 21906767, 21906769, 264681, 33657182, 18108370, 18108374, 18108385, 22279002	264259, 29331822, 29331826, 264905, 264906, 264908, 264510, 265009, 264595, 264758, 265011, 87168559, 265017, 265018, 265019, 264448, 264766, 264686, 21906767, 21906767, 21906769, 265020, 265021, 26176015, 264690, 264692, 264693, 18108368, 18108370, 263972, 55810764, 22479000, 22279000	66714117, 264508, 264509, 264906, 264907, 264908, 264511, 264910, 264764, 264687, 264689, 33657109, 35696423, 35695855, 264632	29331825, 33109954, 264369, 264767, 264689, 33657109, 83373044	264259, 29331824, 29331827, 60433438, 265022, 264636	35696286, 22278997, 29331825, 264909, 21906754, 265017, 265018, 265019, 264682, 264683, 264688, 21906766, 21906767, 21906768, 21906769, 264691, 264555, 264556, 22279000, 264566	264906, 264909, 264632, 18108381
immunoglob	helicase	transport	kinase	UNCLASSIFIED	UNCLASSIFIED	kinase	deaminase
Contains protein domain (PF00047) - immunoglob Immunoglobulin domain	Contains protein domain (PF00575) - helicase S1 RNA binding domain	Contains protein domain (PF01412) - transport Putative GTP-ase activating protein for Arf				Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	
Novel Protein sim. GBank gi 1083506 pir S50065 - sialoadhesin - mouse	Novel Protein sim. GBank gij3913431(sp 042643 DDX8_SCHPO - PUTATIVE PRE- MRNA SPLICING FACTOR ATP-DEPENDENT RNA HELICASE SPAC10F6.02C	Novel Protein sim. GBank gi 5689535 dbj BAA83051.1 - (AB029022) KIAA1099 protein [Homo sapiens]	Novel Protein sim. GBank gi 2076894 gb AAB53983.1 - (AF002197) short region of weak similarity to protein kinase C; contains similarity to Pfam domain PF00130 (DAG_PE-bind), Score=10.0, E-value=0.0034, N=1 [Caenorhabditis elegans]	Novel Protein sim. GBank gi 3875371 emb CAA85414.1 - (Z36948) contains a valine and arginine rich domain, possesses weak similarity with the RNA binding domains from RNA splicing factor U2AF 65 KD subunit; cDNA EST EMBL:D64658 comes from this gene; cDNA EST EMBL:D66829 comes fr			Novel Protein sim. GBank gi 1814270 (U74586) - double- stranded RNA specific adenosine deaminase [Rattus norvegicus]
1665 91226952 (3329, 3330) Novel Protein sim. sialoadhesin - mou	95358160 (3331, 3332) Novel Protein sim. 9i[3913431[sp]042 MRNA SPLICING HELICASE SPAC	91228655 (3333, 3334) Novel Protein sim. (AB029022) KIAA	88095135 (3335, 3336) Novel Protein sim. (AF002197) short (C; contains similar bind), Score=10.0, elegans)	91227846 (3337, 3338)	87628009 (3339, 3340)	87346372 (3341, 3342) Novel Protein sim. gil462451 sp P342 SERINE/THREON	1672 86291834 (3343, 3344) Novel Protein sim. stranded RNA spen norvegicus]
1665	1666	1667				1671	1672

264488, 264569, 18108394, 56994075, 22278996, 264293, 35695652, 264508, 264905, 264509, 264906, 264907, 264908, 264908, 264509, 264906, 264907, 264908, 264594, 264595, 264700, 60170831, 264592, 264594, 264583, 264764, 264288, 264601, 264766, 264688, 264768, 264687, 264769, 264689, 264689, 264690, 3055709, 264628, 264691, 264633, 3569423, 3569585, 264631, 264632, 264634, 264634, 264634, 264637, 264637, 264638, 264568, 264567, 264486	29331822, 29331824, 264906, 52644045, 60433356, 87168559, 264448, 264288, 264688, 264691				SIFIED 29331824, 264102		65274572, 22278994, 22278999, 66714117, 29331827, 56182435, 21906754, 265018, 264288, 21906769		SIFIED 264909, 265017, 264605
Kinase	kinase	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	carboxyli	transport	UNCLASSIFIED	UNCLASSIFIED
Contains protein domain (PF00130) - kinase Phorbol esters/diacy/glycerol binding domain (C1 domain)	_					Contains protein domain (PF00289) - carboxylase Carbamoyl-phosphate synthase (CPSase)			1
				Novel Protein sim. GBank gil3327046 dbj BAA31591 - (AB014516) KIAA0616 protein [Homo sapiens]		91214106 (3357, 3358) Novel Protein sim. GBank gi 550452 (U08469) - 3- methylcrotonyl-CoA carboxylase, biotin-carrier domain [Glycine max]	91005372 (3359, 3360) Novel Protein sim. GBank gi 2394478 (AF024500) - No definition line found [Caenorhabditis elegans]	Novel Protein sim. GBank gi 5689537 db BAA83052.1 - (AB029023) KIAA1100 protein [Homo sapiens]	
88095137 (3345, 3346)	1674 88258028 (3347, 3348) Novel Protein sim. (AL080062) hypott	87606466 (3349, 3350)			86866829 (3355, 3356)	91214106 (3357, 3358)	91005372 (3359, 3360)		1682 86042710 (3363, 3364)
1673				1677				1681	1682

1683	1683 94316213 (3365, 3366) Novel Protein sim. Gi gi 5031717 reflNP_00	Novel Protein sim. GBank gil5031717 ref NP_005704.1 pGPBP - goodpasture antigen- START domain binding protein	Contains protein domain (PF01852) - UNCLASSIFIED START domain	UNCLASSIFIED	263994, 35696286, 35696052, 264508, 264509, 264509, 264905, 264907, 264907, 264908, 264907, 264908, 264907, 264908, 264909, 264909, 264910, 264511, 265006, 265010, 265011, 264761, 264763, 264763, 264764, 264369, 264768, 35695917, 264692, 33657109, 35696423, 35695855, 264634, 264635, 264636, 264638, 264639, 83373044, 264486
1684	80063409 (3367, 3368)			UNCLASSIFIED	264563, 264566
1685	94323182 (3369, 3370) Novel Protein sim. Glby C. elegans cDNA cDNA yk34a9.3; Sim [Caenorhabditis eleg	Novel Protein sim. GBank gij1255371 (U53147) - coded for by C. elegans cDNA yk34a9.5; coded for by C. elegans cDNA yk34a9.3; Similar to guanylate kinase. [Caenorhabditis elegans]	Contains protein domain (PF00625) - Guanylate kinase	kinase	60424179, 52646842, 22278994, 35696286, 22278998, 264259, 52645080, 29331824, 28931826, 265007, 33657084, 265018, 264881, 26448, 264683, 264689, 21906765, 21906767, 21906768, 21906769, 265021, 264692, 65274620, 33657109, 22279000
1686	87820710 (3371, 3372)	1686 87820710 (3371, 3372) Novel Protein sim. GBank gi 2244707 dbj BAA21115.1 - (AB005287) thrombospondin 1 [Bos taurus]		UNCLASSIFIED	56182575, 264259, 264508, 264905, 264509, 264907, 264908, 264510, 264511, 265006, 264512, 265006, 264512, 265009, 264511, 265006, 265011, 264605, 18108351, 264764, 264766, 18108357, 264638, 264631, 264631, 264635, 264635, 264635, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264638, 264638, 264638, 264638, 264638, 264638, 264567
1687	94719400 (3373, 3374)	94719400 (3373, 3374) Novel Protein sim. GBank gil4680679lgblAAD27729.1 AF13295 - (AF132954) CGI-20 protein [Homo sapiens]		UNCLASSIFIED	35695917, 264906, 264907, 264908, 264510, 265006, 265007, 264910, 264556, 18108381, 18108383, 265011
1688	82158442 (3375, 3376)			UNCLASSIFIED	264905, 264910, 264760, 264629, 264555
1689	94325049 (3377, 3378)	94325049 (3377, 3378) Novel Protein sim. GBank gi 4240193 db BAA74875.1 - (AB020659) KIAA0852 protein [Homo sapiens]		UNCLASSIFIED	264569, 35696286, 22278999, 264508, 264908, 264909, 56182435, 33657402, 60433438, 55812038, 265017, 265018, 264448, 264764, 264288, 264686, 29148629, 35695917, 265020, 265021, 263972, 18108374, 65274791, 83373044, 264089
1690	83255346 (3379, 3380)	83255346 (3379, 3360) Novel Protein sim. GBank gij3800736 (AF031572) - seven- pass transmembrane receptor precursor [Mus musculus]	Contains protein domain (PF00028) - cadherin Cadherin	cadherin	264369, 21906766, 264692, 264639, 87168518
1691	88095223 (3381, 3382)				264768, 33657109, 29331827, 29148629, 264510, 264106, 264910, 264109, 264508, 60170831, 264563, 264905, 264564, 264691, 264637, 264628, 264907, 264908, 33657023, 264567, 264766, 263974
1692	86106709 (3383, 3384)				264106
1693	87012775 (3385, 3386)	87012775 (3385, 3386) Novel Protein sim. GBank gi 121271 sp P02207 GLB_LAMFL - GLOBIN	Contains protein domain (PF00042) - UNCLASSIFIED Globin	UNCLASSIFIED	29331826, 264508, 264905, 264907, 264595, 265010, 265011, 21906768, 33657023, 264529, 263978, 264558

35596286, 22278996, 22278998, 22278999, 264259, 29331822, 29331825, 29331826, 29331826, 29331826, 29331826, 29331826, 29331828, 35596052, 29146499, 264905, 264908, 52644045, 264511, 265006, 265007, 265009, 264592, 60433356, 21906754, 265010, 265011, 18108351, 264763, 264689, 21906768, 60170615, 264691, 264692, 264693, 18108370, 18108374, 263978, 35696423, 3569555, 26456, 18108381, 18108388, 87168518, 264482, 264488	UNCLASSIFIED 284634	22278997, 264259, 265010, 18108351, 264764, 21906766, 18108370	glycoprotein 264682	56994075, 22278998, 22278999, 264509, 33657402, 264758, 87168474, 87168559, 265017, 265018, 264448, 264687, 29148627, 21906769, 29148629, 265020, 265022, 33657023, 264558, 87168518, 22279002	UNCLASSIFIED 35696286, 264635		UNCLASSIFIED 264092, 264110, 263977		259115, 255017, 264762, 18108351, 264288, 21905755, 3695917, 255017, 255021, 60170615.	263967, 33657109, 18108370, 263972, 263974, 18108374, 263976, 35695855, 264555, 263881, 60170394, 18108385, 56526486, 87168518, 60432113	UNCLASSIFIED 264909, 265017, 264628, 264629, 264638	UNCLASSIFIED 264369	264569, 35696286, 264907, 265010, 264687, 264693, 264696, 264693, 264698, 264699, 264690, 264699, 264690, 264690, 264690, 264690, 264690, 264690, 264690, 264690, 264690, 264690, 264690, 264690, 264690, 264690, 264690, 2646	22278996, 22278998, 264259, 264509, 265018, 264764, 264685, 284686, 21906788, 21906789, 265022, 264691, 264558, 22279000
Contains protein domain (PF0193) - rnapolymerase RNA polymerases L / 13 to 16 kDa subunit			giyco		ONO	Contains protein domain (PF00096) - nucl_recpt Zinc finger, C2H2 type		MHC				ONO		
Novel Protein sim. GBank gi[\$453932 ref NP_006225.1 pPOLR - polymerase (RNA) II (DNA directed) polypeptide J (13.3kD)	94719325 (3389, 3390) Novel Protein sim. GBank gil4680679lgb AAD27729.1 AF13295 - (AF132954) CGI-20 protein [Homo sapiens]	Novel Protein sim. GBank gil4220517 emb CAA22990 - (AL035356) hypothetical protein [Arabidopsis thaliana]	85740963 (3393, 3394) Novel Protein sim. GBank gij505652 (U10362) - GP36b glycoprotein [Homo sapiens]	87445285 (3395, 3396) Novel Protein sim. GBank gi[5052031]gb AAD38411.1 AF15573 - (AF155739) axotrophin [Mus musculus]		87859161 (3399, 3400) Novel Protein sim. GBank gi[543344 pir S41647 - zinc finger 5 protein - mouse		87795092 (3403, 3404) Novel Protein sim. GBank gij3877439 emb CAA96652 - (272510) similarity to yeast UTR3 protein (Swiss Prot accession number P21374) cDNA EST EMBL (17282)	comes from this gene; cDNA EST EMBL:D75763 comes from this gene; cDNA EST yk274e3.3 comes from this	gene; cDNA EST yk274e3.5 c	79568651 (3405, 3406) Novel Protein sim. GBank gil451544 (U04267) - proline-rich cell wall protein [Gossypium barbadense]	86622979 (3407, 3408) Novel Protein sim. GBank gi 1263289 (U47856) - fibroin-4 [Araneus diadematus]	87795175 (3409, 3410) Novel Protein sim. GBank gil4519621[dbjjBAA75670.1] - (AB017614) OASIS protein [Mus musculus]	87790967 (3411, 3412) Novel Protein sim. GBank gi[3123034[sp Q15011 Y025_HUMAN - HYPOTHETICAL PROTEIN KIAA0025
94208168 (3387, 3388)	94719325 (3389, 3390)			87445285 (3395, 3396)	1699 87424793 (3397, 3398)	1700 87859161 (3399, 3400)	1701 86570488 (3401, 3402)	2 87795092 (3403, 3404)					-	1706 87790967 (3411, 3412)

264569, 264259, 29331825, 29331826, 29331826, 25931826, 3569602, 264509, 264905, 264907, 264900, 264512, 265009, 264910, 264592, 264595, 264758, 265017, 264681, 264764, 264766, 264686, 18108357, 35695917, 264690, 264693, 264633, 264638, 264639, 18108388, 18108380, 264638, 264639, 18108388, 18108391, 264638, 265011, 264681, 264684, 265011, 264681, 264684, 265011, 264681, 264684, 265011, 264681, 264684, 265011, 264681, 264684, 265011, 264681, 265011, 264681, 264684, 265011, 264681, 264684, 265011, 264681, 264684, 265011, 264681, 265011, 264681, 264684, 264684, 265011, 264681, 264684, 264684, 264681, 265011, 264681, 264684, 264681, 264681, 265011, 264681, 264684, 264681, 2	2030 1, 204031, 204062, 2040604, 2040600, 204689, 21906765, 265021, 264691, 33657023, 264693, 18108370, 35695855, 26279002, 22279002	10100390, 032/43/1, 30080200, 222/18991, 60432209, 222/302049, 56182181, 66714117, 60432289, 604322049, 36502049, 36696052, 29331828, 264596, 60433438, 55812038, 264759, 265010, 264601, 265017, 264448, 264764, 264288, 264769, 21906766, 21906769, 55811957, 265020, 265021, 52644150, 33657109, 33657182, 27486262, 33657349, 35695763, 18108374, 35696423, 56811676, 3663685, 2646151, 56182323, 264559, 264564, 264486	56182575, 22278999, 264259, 29331824, 66432289, 29331827, 35696052, 264508, 264905, 264907, 264907, 264907, 264907, 264908, 264510, 264510, 264758, 21906754, 265011, 264681, 18108357, 264689, 21906755, 55811957, 264693, 20281149, 264629, 18108374, 55811376, 65274791, 264630, 20281071, 264634, 264635, 264636, 264637, 264556, 264638, 264639, 56182323, 87168518	22278994, 22278999, 29331822, 265006, 265007, 265008, 55812038, 21906754, 60174639, 265011, 87168559, 18108351, 118108354, 21906765, 21906766, 21906766, 21906769, 265020, 33657109, 18108374, 264556, 60170394, 83373044, 18108384, 264486	56994075, 29331824, 29331828, 265009, 18108351, 21906768, 265020, 33657023, 18108374, 83373044
UNCLASSIFIED	UNCI ASSIFIED		UNCLASSIFIED	UNCLASSIFIED	kinase
					Contains protein domain (PF00567) - kinase Tudor domain
1717 87563223 (3433, 3434) Novel Protein sim. GBank gil2765411emb CAA74749 - (Y14391) GTP-binding protein [Homo saplens] (Y14391) GTP-binding protein [Homo saplens] 1718 87032628 (3435, 3436) Novel Protein sim. GBank	gil2833262 sp 014999 Y076_HUMAN - HYPOTHETICAL PROTEIN KIAA0076 (HA0936) 94315259 (3437, 3438) Novel Protein sim. GBank	gil4505197[ref]NP_003473.1[pMLL2 - myeloid/lymphoid or mixed-lineage leukemia 2	e4033003 (3439, 3440), Novel Protein sm. GBank gij2/294/8/pir/[S51939 - chitinase (EC 3.2.1.14) precursor - beet	91722288 (3441, 3442) Novel Protein sim. GBank gij4886461 emb CAB43381.1 - (AL050280) hypothetical protein [Homo sapiens]	94134549 (3443, 3444) Novel Protein sim. GBank gij5689375 dbj BAA82968.1 - (AB030644) tudor repeat associator with PCTAIRE 2 (Rattus norvegicus)
1717 87563223 (3433, 3434) 1718 87032628 (3435, 3436)	1719 94315259 (3437, 3438)			1721 91722288 (3441, 3442)	1722 94134549 (3443, 3444)

264488, 264687, 264769, 21906767, 21906768, 56182575, 55811957, 22278997, 22278998, 265020, 264259, 264692, 33657023, 29331824, 29331825, 60432289, 33657182, 33656970, 33657349, 264629, 264908, 264907, 18108370, 264629, 264908, 264907, 18108370, 264910, 264501, 265006, 264511, 265007, 264910, 264532, 264591, 60432229, 264592, 60433356, 264594, 60433438, 264595, 83373044, 55812038, 33109954, 33557084, 87168518, 87168474, 265010, 265011, 87168592, 264600, 60432113, 264604, 265019, 264567, 264488, 264682, 264369, 264766, 264488, 264488, 264486, 264369, 264766	264905, 18108359, 264693, 264628, 264631, 264636, 264555, 264556, 264558, 264559	35696286, 264269, 29331822, 35696052, 264508, 264509, 264906, 264906, 264907, 264908, 264909, 265910, 26509, 264591, 264801, 264760, 18108351, 264881, 264766, 264768, 21906769, 35695917, 264628, 35696423, 264639, 264631, 264632, 264635, 264638, 264656, 264536, 264638, 264656, 264636, 264658, 2	29146498, 264683, 264689	264905, 265011, 264689, 21906768
UNCLASSIFIED	UNCLASSIFIED	homeobox	UNCLASSIFIED	UNCLASSIFIED
	Contains protein domain (PF00096) - Zinc finger, C2H2 type	Contains protein domain (PF00023) - homeobox Ank repeat		
Novei Protein sim. GBank gi[4426962]gb AAD20633] - (AF126062) Arf-like 2 binding protein BART1 [Homo sapiens]	GBank gi 2340162 (AF005083) - dsRBP- vis]	GBank gi 3152662 (AF064604) - KE03 iens]		Novel Protein sim. GBank gil4689348 gb AAD27861.1 AF13256 - (AF132562) BcDNA_LD14270 [Drosophila melanogaster]
1723 95358181 (3445, 3446) Novel Protein sim. (AF126062) Arf-like sapiens]	1724 87713806 (3447, 3448) Novel Protein sim ZFa (Xenopus lae	85655191 (3449, 3450)	85754255 (3451, 3452)	1727 85296362 (3453, 3454) Novel Protein sim- gil4689348 gb AAI BcDNA.LD14270
1723	1724	1725	1726	1727

		UNCLASSIFIED 29331822, 29331825, 29331828, 264907, 264908, 264909, 265011, 264764, 264629	UNCLASSIFIED 52646842, 264907, 264909, 56182435, 55811386, 87168559, 265018, 265019, 264760, 52644229, 55811576	struct 56182575, 29331822, 29331824, 29331827, 66712502, 264591, 33657402, 60433356, 265019, 21906768, 21906769, 35695917, 265020, 265021, 264636, 56182323	3 - Synthase 264907, 264910, 33657402, 265010, 264681, 264684, 264686, 264769, 264691, 264693, 264693, 264638, 264638, 264636, 264556	264693
					Contains protein domain (PF00640) - synthase Phosphotyrosine interaction domain (PTB/PID).	
1728 95349515 (3455, 3456) Novel Protein sim. GBank gil4406549lgb[AAD20027[-	Novel Protein sim. GBank gil854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6]		Novel Protein sim. GBank gij631600 pir S47094 - hypothetical protein - rabbit	91218878 (3463, 3464) Novel Protein sim. GBank gij4240231 dbj BAA74894.1 - (AB020678) KIAA0871 protein [Homo sapiens]	Novel Protein sim. GBank gil1575756 (U70674) - m-Numb [Mus musculus]	
95349515 (3455, 3456) [91227948 (3457, 3458) Novel Protein sim. G (X83413) U88 [Hum	85483474 (3459, 3460)	88266068 (3461, 3462) Novel Protein sim. G hypothetical protein	91218878 (3463, 3464)	87617178 (3465, 3466) Novel Protein sim. [Mus musculus]	1734 87795261 (3467, 3468)
1728	1729	1730	1731	1732	1733	1734

264488, 18108394, 18108398, 22278998, 60432049, 264259, 29331822, 264908, 265006, 265007, 265008, 265009, 264591, 33657402, 265017, 8716859, 265077, 18108351, 264682, 18108354, 265077, 18108351, 264682, 21906765, 21906768, 21906769, 29148629, 29148784, 265021, 265022, 25644150, 18108364, 18108365, 33657109, 18108374, 18108380, 18108385, 87168518, 264553, 18108380,	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33657023, 33657109, 264555, 60431850, 264637, 60170394, 264558, 264639, 264564	29331822, 29331830, 264591, 265011, 265018, 265019, 22279002	264908, 264909, 265008, 264910, 264566	35696052, 264603, 264557	264604, 21906764, 18108364, 264629, 35695855, 264636	60432289, 29331827, 264509, 265009, 60432229, 264759, 265017, 264767, 264688, 264689, 21906769, 265020, 33657109	264906, 264910, 264758, 265011, 264631, 264638, 26456	265017, 265020, 264692	22278998, 264508, 264907	264558	65274791, 264639, 264559	52646842, 29331824, 29331825, 52644045, 56182435, 265007, 52646317, 87168474, 265018, 18108351, 264369, 264769, 264689, 35695917, 60170615, 264691, 33657023, 65274791, 35695855, 60170394, 22279000, 264482
	UNCLASSIFIED	kinase	UNCLASSIFIED	UNCLASSIFIED	traffic	UNCLASSIFIED	UNCLASSIFIED		UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	transcriptfactor
Contains protein domain (PF01169) - Uncharacterized protein family UPF0016		Contains protein domain (PF00168) - kinase C2 domain		Contains protein domain (PF00441) - UNCLASSIFIED Acyl-CoA dehydrogenase		Contains protein domain (PF01363) - UNCLASSIFIED FYVE zinc finger					Contains protein domain (PF00400) - UNCLASSIFIED WD domain, G-beta repeat	
1 AE14679 - (AE146793) PET27	95362884 (3471, 3472) Novel Protein sim. GBank gil4885647 refNP_005472.1 pTRAP - thyroid hormone receptor-associated protein complex component	88165549 (3473, 3474) Novel Protein sim. GBank gil2143607 pir S68695 - B/K protein - rat	85788811 (3475, 3476) Novel Protein sim. GBank gi 2225941 emb CAA69714 · (Y08460) Mdes protein [Mus musculus]		83592939 (3479, 3480) Novel Protein sim. GBank gil4809 emb CAA44309 - X62452} YCR601 [Saccharomyces cerevisiae]	95010100 (3481, 3482) Novel Protein sim. GBank gjl4883898lgb AAD31695.1 AF13042 - (AF130420) serine protease-like protein isoform [Homo sapiens]	85788814 (3483, 3484) Novel Protein sim. GBank gil4505193 ref NP_003667.1 pMLD - membrane fatty acid (lipid) desaturase		91224003 (3487, 3488) Novel Protein sim. GBank gij728832 sp P39189 ALU2_HUMAN - !!!! ALU SUBFAMILY SB WARNING ENTRY !!!!		94326110 (3491, 3492) Novel Protein sim. GBank 9i 731756 sp P38873 YHY6_YEAST - HYPOTHETICAL 175.8 KD PROTEIN IN GND1-IK11 INTERGENIC REGION	94324333 (3493, 3494) Novel Protein sim. GBank gij 1658503 (U75467) - Atu [Drosophila melanogaster]
1735 88318638 (3469, 3470) Novel Protein sim. GBank gil4836807[gb AAD30566. [Mus musculus]	95362884 (3471, 3472)			87328576 (3477, 3478)	_	95010100 (3481, 3482)	85788814 (3483, 3484)	86966475 (3485, 3486)	91224003 (3487, 3488)	20290075 (3489, 3490)	94326110 (3491, 3492)	94324333 (3493, 3494)
1735	1736	1737	1738	1739	1740	1741	1742	1743	1744	17	1746	1747

1748	1748 88003580 (3495, 3496) Novel Protein sim. GB gil4504511[ref]NP_001 DNAJ-like 2	Novel Protein sim. GBank gil4504511fref NP_001530.1 pHSJ2 - heat shock protein, DNAJ-like 2	Contains protein domain (PF00684) - eph DnaJ central domain (4 repeats)	ebh	264489, 56182575, 29331824, 56182435, 264112, 265007, 265019, 264764, 21906768, 265020, 264691, 55811576, 264635, 264555, 264556, 264557, 264559
1749		83363091 (3497, 3498) Novel Protein sim. GBank gij5650780[gb AAD45948.1 AF15196 - (AF151968) RGS protein RGS-17 [Gallus gallus]	Contains protein domain (PF00615) - oncogene Regulator of G protein signaling domain	опсодепе	264106
1750	94321664 (3499, 3500)	94321664 (3499, 3500) Novel Protein sim. GBank gil4996894 gb AAC28444.2 - (AF065164) hyperpolarization-activated, cyclic nucleotide-gated channel 2 [Homo sapiens]			33657402, 264288, 52644150, 263974, 83373044
1751	83373058 (3501, 3502)	83373058 (3501, 3502) Novel Protein sim. GBank gil2760161 dbj BAAZ4184 - (AB010054) outer arm dynein light chain 2 [Anthocidaris crassispina]	Contains protein domain (PF00560) - ATPase_associated 265010, 264369 Leucine Rich Repeat	ATPase_associated	265010, 264369
1752		86456530 (3503, 3504) Novel Protein sim. GBank gi]3915482[sp P74346]YG29_SYNY3 - HYPOTHETICAL 36.0 KD PROTEIN SLR1629	Contains protein domain (PF00849) - deaminase RNA pseudouridylate synthase	deaminase	264510, 264593, 264682, 21906765, 18108370
1753		94235159 (3505, 3506) Novel Protein sim. GBank gi[2852636 (AF007155) - unknown [Homo sapiens]	Contains protein domain (PF01553) - phosphatase Acyltransferase	phosphatase	56994075, 22278996, 264908, 60170831, 264682, 264764, 264369, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 35695855, 264637, 264564
1754		Novel Protein sim. GBank gif31421 sp P39981 YEH4_YEAST - HYPOTHETICAL 53.3 KD PROTEIN IN HXT8-CAN1 INTERGENIC REGION		transport	264488, 35696286, 264509, 264906, 264907, 264908, 264909, 264511, 264910, 264591, 33657402, 264594, 264757, 264758, 264600, 264604, 264762, 264763, 264685, 264634, 264638, 35696423, 264632, 264634, 264584, 264564, 264565, 264565, 264565
1755		Novel Protein sim. GBank gi 1176422 (U43194) - rhophilin [Mus musculus]		UNCLASSIFIED	264686
1756	92962614 (3511, 3512)	92962614 (3511, 3512) Novel Protein sim. GBank gi 4432860 gb AAD20708 - (AC006300) putative glucose-induced repressor protein [Arabidopsis thaliana]			52646842, 22278994, 22278995, 56994075, 22278996, 22278997, 264259, 29331822, 60432289, 29331822, 255009, 60432289, 29331827, 33656970, 265006, 265009, 6043229, 60433356, 60433438, 33109954, 21906754, 265017, 265018, 265019, 26448, 265017, 265018, 2906766, 21906767, 21906768, 21906769, 248620, 266021, 266022, 264692, 27486262, 2486264, 18108376, 20281152, 264558, 22279002, 264482
1757	95357380 (3513, 3514)	95357380 (3513, 3514) Novel Protein sim. GBank gij5441615 emb CAB46856.1 - (AJ388557) zinc finger protein [Canis familiaris]	Contains protein domain (PF00096) - dna_ma_bind Zinc finger, C2H2 type	dna_rna_bind	22278997, 264259, 60432289, 29331827, 264908, 5264045, 265008, 264593, 265019, 264766, 21906768, 65274620, 18108385, 60432113, 264566, 264487

22278994, 22278999, 29331822, 29331824, 29331825, 33656970, 264508, 265006, 265007, 265007, 265007, 265007, 265017, 265018, 2906769, 265020, 265021, 33657023, 33657109, 264629, 18108374, 35695655, 264632, 52844332, 22279002, 264563	264759	56182575, 60432049, 35696052, 264905, 264906, 264907, 264908, 264909, 265006. 265009, 264909, 265006. 265009, 2649010, 6043229, 264592, 264595, 26412038, 264766, 264769, 264769, 21906765, 25811957, 35695917, 264690, 264692, 264632, 264634, 264636, 264486, 264565, 264639, 2646313, 264565, 264486	29331822, 264910, 264685, 264686	22278999, 264259, 264905, 264907, 264448, 60170831, 265010, 265011, 265017, 264448, 21906765, 21906766, 21906767, 21906768, 265021, 264690, 33657109, 18108374, 264558, 60170394	56181686, 29331825, 35696052, 264905, 264906, 264906, 264909, 264763, 264682, 264769. 35695917, 265022, 33657023, 18108374, 35696423, 264634	22278998, 264259, 35696052, 29331828, 56182435, 265008, 265017, 265018, 264448, 284288, 21906766, 21906767, 29148627, 35695917, 264691, 33657023, 60432113, 22279002	264488, 18108394, 22278995, 22278997, 22278999, 264259, 264104, 264508, 264905, 265009, 265009, 265009, 265009, 265009, 265009, 265009, 265019, 18108351, 264682, 264369, 21906765, 29148627, 29148629, 35695917, 265020, 265021, 264692, 264628, 264629, 18108374, 263976, 264636, 83373044, 22279000, 22279000
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	desaturase		collagen	UNCLASSIFIED
			Contains protein domain (PF00618) - Guanine nucleotide exchange factor for Ras-like GTPases; N-terminal motif	Contains protein domain (PF00173) - desaturase Heme-binding domain in cytochrome b5 and oxidoreductases		Contains protein domain (PF01410) - Fibrillar collagen C-terminal domain	
1758 87612971 (3515, 3516) Novel Protein sim. GBank gij3881040 emb CAA16403 - (AL021497) predicted using Genefinder (Caenorhabditis elegans]		87329716 (3519, 3520) Novel Protein sim. GBank gil5262748 emb CAB45688.1 - (AJ133120) Proline rich synapse associated protein 2 [Rattus norvegicus]) Novel Protein sim. GBank gij127749 sp P10569 MYSC_ACACA - MYOSIN IC HEAVY CHAIN	95319887 (3523, 3524) Novel Protein sim. GBank gil3169158 (AC004770) - BC269730_2 [Homo sapiens]	1763 91224013 (3525, 3526) Novel Protein sim. GBank gil4809026 gb AAD30062.1 - (AF132856) suppressor of G2 allele of skp1 homotog [Homo sapiens]	87757697 (3527, 3528) Novel Protein sim. GBank gij 1360869 pir CGHU1V - collagen alpha 1(V) chain precursor - human	91230091 (3529, 3530) Novel Protein sim. GBank gij486806lpirij535503 - finger protein neuralized - fruit fly (Drosophila melanogaster)
58 87612971 (3515, 3516)	59 36994372 (3517, 3518)	1760 8 7329716 (3519, 3520)	1761 87409586 (3521, 3522) Novel Protein sim. G gil127749 sp P10566 CHAIN	1762 95319887 (3523, 3524)	763 - 9,7224013 (3525, 3526	1784 87757697 (3527, 3528	1765 91230091 (3529, 3530

									_		_							_		_			
52645156, 87168559, 60170615, 33657023, 264693, 33657109, 27486261, 264555, 83373044	265017, 265019, 264686, 264768, 265020, 264692	29331824, 29331825, 264591, 56182323	264563	264488, 264768, 264769, 56182575, 55811957, 264690, 264691, 35696052,	264905, 264509, 264906, 264907, 264628, 264908, 264908, 264900, 264634, 264635, 26456, 264757, 264758, 55812038,	65274444, 264760, 264563, 264762, 264764, 264684, 264766	264758, 264600, 264369, 55811957, 265020,	833/3044, 222/9000	22278998, 29331828, 33109954, 265018,	265019, 264764, 21906765, 265020, 265021, 264556	65274572, 56182575, 22278996, 35696286, 22278999, 35696286, 22278999	29331825, 60432289, 35696052, 264106,	264509, 264906, 264907, 29331830, 264908,	52644045, 264511, 265006, 265007, 265006, 60170831, 60433438, 264758, 55811386.	87168559, 265017, 264604, 265019,	55811150, 264288, 56181562, 264689,	21906/66, 21906/67, 21906/68, 21906/69, 65811067, 265020, 265020, 52644150	[264691, 33657023, 264692, 264693,	60431528, 35696423, 35695855, 264636,	56182323, 18108387, 56526486, 22279000,	22279002, 264563, 264564, 264565, 264566, 264567	65274572, 56182575, 35696052, 55812038, 33109954, 21906754, 265017, 21906767,	21906768, 21906769, 265020, 264691, 264636, 56182323, 22279002
glycoprotein	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	complement			transferase				potassium_channel											UNCLASSIFIED	
				Contains protein domain (PF00089) - complement Trypsin			Contains protein domain (PF00515) - transferase	TPR Domain	Contains protein domain (PF01529) -	DHHC zinc finger domain	Contains protein domain (PF01412) - potassium_channel	Furative G i P-ase activating protein for Arf										Contains protein domain (PF00415) - UNCLASSIFIED Requistor of chromosome	condensation (RCC1)
1766 95081201 (3531, 3532) Novel Protein sim. GBank gijz499087lsp Q09332 UGGG_DROME - UDP. GLUCOSE:GLYCOPROTEIN GLUCOSYLTRANSFERASE PRECURSOR (DUGT)	87755998 (3533, 3534) Novel Protein sim. GBank gi 4176443 emb CAA18263.1 - AL022238) dJ1042K10.4 (novel protein) [Homo sapiens]	3536)	3538)	_	COMPONENT PRECURSOR		94233542 (3541, 3542) Novel Protein sim. GBank	gij39141911sp P56556 OGT1_RAT - UDP-N-ACETYLGLUCOSAMINE-PEPTIDE N-ACETYLGLUCOSAMINYLTRANSFERASE 110 KD SUBUNIT (O-G! CNAC TRANSFERASE P110 SUBUNIT)	1772 87643510 (3543, 3544) Novel Protein sim. GBank	gila959442lgb[AAD34351.1]AF12136 - (AF121360) DNZDHHC/NEW1 zinc finger protein 11 [Drosophila melanooaster]	94116824 (3545, 3546) Novel Protein sim. GBank gil3978464 (AF085693) - G	protein-coupled receptor kinase-associated AUP ribosylation factor GTPase-activating protein [Rattus	norvegicus]									94232573 (3547, 3548) Novel Protein sim. GBank ail2455699isniO15034 Y032_HUMAN - HYPOTHETICAL	PROTEIN KIAA0032
95081201 (3531, 35			87388988 (3537, 3538)	95413144 (3539, 3					97643510 (3543, 3)													4 94232573 (3547, 3	
1766	1767	1768	1769	1770			1771	_	1772		1773											1774	

1775	95359330 (3549, 3550)	95359330 (3549, 3550) Novel Protein sim. GBank gil 1469199 dbj BAA09487 - (D50928) The KIAA0138 gene product is novel. [Homo sapiens]		UNCLASSIFIED	65274572, 56182575, 56994075, 35696286, 22278997, 29331822, 29331826, 60432289, 29331828, 35696052, 29331826, 60432289, 29331828, 35696052, 29331830, 66712602, 264828, 56182435, 264511, 265007, 265009, 60432438, 55812038, 251906754, 85658542, 87168559, 264601, 265017, 265018, 265019, 264762, 26448, 264288, 204689, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 264691, 33657023, 264692, 33657109, 27486261, 33657349, 18108370, 18108377, 35696423, 55811576, 35695855, 264632, 264634, 264636, 264639, 56182323, 83373044, 56526486, 87168518, 60432113, 22279000, 222790002, 264482, 264634, 264634, 264632, 264482, 264482, 264634, 26463
1776	94133756 (3551, 3552)	94133756 (3551, 3552) Novel Protein sim. GBank gil4589676 dbj BAA76857.1 - [AB023230) KIAA1013 protein [Homo sapiens]			264910
1777	87447171 (3553, 3554)		_	nucl_recpt	56994075, 29331826, 265008, 87168474, 265017, 265018, 264761, 55811150, 264764, 56181562, 264689, 21906765, 21906768, 21906769, 35695917, 264690, 33657023, 35695763, 60431528, 35696423, 55811576, 35695855, 22279000, 22279002, 264564
1778		Novel Protein sim. GBank gil3875648 emb CAA91454.1 - (Z66561) Similarity to Human rab13 protein (PIR Acc. No. A49647). Contains the ATP/GTP-binding site motif (PROSITE PS00017); cDNA EST EMBL.:M89412 comes from this gene; cDNA EST yk21299.3 comes from this gene; cDNA EST yk21299		UNCLASSIFIED	29331826, 29331827, 35696052, 264512, 265007, 265009, 265017, 265019, 264762, 18108351, 264769, 21906765, 21906766, 21906767, 21906767, 21906768, 21906769, 265020, 265021, 264691, 264693, 18108370, 264556, 83373044
1779		94133758 (3557, 3558) Novel Protein sim. GBank gil4589676 dbj BAA76857.1 - (AB023230) KIAA1013 protein [Homo sapiens]		UNCLASSIFIED	29148627, 35696286, 29147620, 265006, 265007, 265008, 18108385, 65274727, 264482, 264369, 264766
1780	87023497 (3559, 3560)		Contains protein domain (PF00807) - UNCLASSIFIED Apidaecin	UNCLASSIFIED	264107, 33657109, 56526486
1781	84047477 (3561, 3562)			UNCLASSIFIED	264508, 264906, 264639
1782	88094607 (3563, 3564) Novel Protein sim. - CORTEXIN	Novel Protein sim. GBank gi 729225 sp P41237 CTXN_RAT - CORTEXIN			264269, 29331822, 264508, 264905, 264906, 264907, 264908, 265007, 265009, 264910, 264591, 264758, 264764, 264288, 264768, 264769, 264639, 264639, 264639, 264653
1783		85717905 (3565, 3566) Novel Protein sim. GBank gi 2257543 dbj BAA21436 - (AB004538) protein arginine N-methyltransferase [Schizosaccharomyces pombe]		interferon	264768

UNCLASSIFIED UNCLASSIFIED	264488, 83373044	001100	264488 83373044	264488, 83373044	264488 29331828 264909 18108351	264488, 29331828, 264909, 18108351,	264288, 265021, 264555, 264636	SEEDENEY SEADINE SEADINE SEADINE	35696052, 264905, 264906, 264907, 26	264909, 265008, 264910, 264758, 265011,	265019, 264764, 264766, 264769, 264628, 264635
И69)	UNCLASSIFIED	0111004	INCIASSIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED					
Contains protein domain (PF0C											
1784 95197093 (3567, 3569) Novel Protein sim. GBank gi 1755049 (U55042) - myosin X Contains protein domain (PF00169) - struct		elegans	diagaila		Name Bratis rim CBack all 2088660 (AE003130)	88094529 (3577, 3578) Novel Protein sim. GBank gi 2088669 (AF003130) -	F55A12.9 gene product [Caenorhabditis elegans]	יר שאונג א שנוים אומחתת לספרוטוומחתווא בובאפווא			
1784 95197093 (3567, 3568) [6] [1785 95357475 (3569, 3570) [786 85296465 (3571, 3572) [787] [787] 87434784 (3573, 3574)			19530 3530, 05500000	91228779 (3575, 3576)		Т			1790 82489734 (3579, 3580)		

97259 (3581, 3582)	1791 95197259 (3581, 3582) Novel Protein sim. GBank gi[2114321 dbj BAA20037 - (D88733) membrane glycoprotein [Equine herpesvirus 1]	Contains protein domain (PF00047) - glycoprotein Immunoglobulin domain		264488, 264686, 264687, 264768, 18108394, 264769, 18108397, 264259, 264691, 264692,
				33557023, 264693, 264509, 264905, 264906, 264628, 264907, 264529, 264908, 264909, 26510, 265006, 264511, 265008, 264630, 265009, 264631, 264932, 264632, 264635, 264635, 264635, 264635, 264635, 264637, 264637, 264637, 264638, 26463
				. 264593, 264638, 18108381, 264639, 264758, . 265010, 265011, 264602, 22279000, 264604, . 264760, 264564, 264681, 264762, 264565,
				264763, 264683, 264566, 264764, 264288, 264684, 264567, 18108354, 18108391,
(3583 3584)	87792690 (3583, 3584) Novel Protein sim GBank gil4337106lpbiAAD180821.	Contains profein domain (PF01585) - UNCLASSIFIED	UNCLASSIFIED	264685, 264766 22278997, 264259, 264508, 265007,
(_	G-patch domain		33657402, 87168559, 264369, 33657023,
				83083635, 20261071, 204335, 10106367, 87168518
7 (3585, 3586)	95337877 (3585, 3586) Novel Protein sim. GBank	Contains protein domain (PF01532) - ATPase_associated		65274572, 22278995, 22278996, 22278997,
	gijos/933 ilgopaaD43504. ijar i4573 - (AF 143732) endoplasmic reticultum alpha-mapposidase [[Homo sapiens]	Giycosyi nyarotase ranniy 47	-	66714117, 60432289, 29331827, 29331828,
				264103, 264105, 29331830, 265007, 264910,
				265009, 60170831, 60433356, 21906754,
				265010, 265017, 265019, 264681, 264682, 264288, 52644330, 24006766, 24006766
				264266, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 265020,
				265021, 265022, 60170615, 52644150,
				33657023, 33657109, 18108370, 18108374,
				65274791, 20281071, 60432113, 22279000,
				264482, 264564
6 (3587, 3588)	87759806 (3587, 3588) Novel Protein sim. GBank gil4914604 emb CAB43677.1 -	Contains protein domain (PF01798) - UNCLASSIFIED	UNCLASSIFIED	18108394, 22278995, 22278999, 264259,
	(AL050369) hypothetical protein [Homo sapiens]	Putative snoRNA binding domain		29331822, 29331824, 29331825, 29146498,
				29146499, 264508, 264905, 52644045,
				264112, 265006, 265008, 264910, 60433356,
				264/57, 55612035, 67156474, 265011, 565017, 18108351, 564683
				264369, 21906765, 21906766, 21906767,
				21906769, 29148784, 35695917, 60170615.
				33657023, 264629, 18108374, 18108376,
				35696423, 35695855, 264556, 264557,
79747856 (3589 3590)	1		UNCLASSIFIED	264632, 264635, 264636, 264595, 264596,
2000				264907, 264566, 264909
36 (3591, 3592)	86599486 (3591, 3592) Novel Protein sim. GBank		glycoprotein	264488, 264907, 264909, 264594, 264595,
				264766, 264687, 21906765, 21906767, 264628, 264630, 264630
	TACTOR 6, WILLOUISING THE CONSON (MEL '9)			

ank gi[2832906]dbig BAA24608.1 - eptidase III [Ratus nonegicus] 2 protein [Homo sapiens] 2 protein [Homo sapiens] 3ank SSSP_VOLCA - SULFATED 30TEIN 185 (SSG 185) 3ank T29.1 AF13295 - (AF132954) CGI-20 3ank T29.1 AF13295 - (AF132954) CGI-20 3ank T29.1 AF13295 - (AF132954) CGI-20 3ank	IArabi	1787 91223719 (3393, 3394) Novel Protein sim. GBank gi 1842111 (U87586) - decoy fArabidonsis thatianal	ribosomalprot	22278996, 22278997, 22278998, 22278999,	
Poptidase eph eph eph eph eph eph eph eph eph ep	l			, 255, 1524, 1575, 15, 15, 15, 15, 15, 15, 15, 15, 15, 1	
Peptidase eph eph eph eph eph eph eph eph eph ep	- 11			33657023, 263967, 33657109, 22279000	
DO CO	٠.	rotein sim. GBank gi[2832906 dbj BAA24608.1 -	peptidase	22278994, 56994075, 22278997, 22278998,	Γ -
DO CLASSIFIED UNCLASSIFIED	₹	opepiidyi pepiidase III (Kaitus norvegicus)		22278999, 264259, 29331826, 60432289,	
UNCLASSIFIED UNCLASSIFIED				264757, 60433438, 21906754, 33657084,	
UNCLASSIFIED UNCLASSIFIED				87168559, 265017, 18108351, 264682,	
20 UNCLASSIFIED				264448, 264288, 21906765, 21906766, 21006767, 21006769	
20 COUNCLASSIFIED				21500101, 21500100, 21500103, 55035311, 265020, 265021, 33657023, 33657182	
UNCLASSIFIED UNCLASSIFIED				27486261, 27486265, 33657349, 263973,	
20 COUNCLASSIFIED				18108374, 55811576, 35695855, 18108385,	
Homo sapiens UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED UNCLA - SULFATED UNCLA - SULFATED UNCLASSIFIED UNCLASSIFIE		Protein sim. GBank gil5689541 dbj BAA83054.1 -	eph	264908, 21906754, 21906767, 21906769,	Т
14F13295 - (AF132954) CGI-20 P_VOLCA - SULFATED SIN 185 (SSG 185) 1 AF13295 - (AF132954) CGI-20	N	9025) KIAA1102 protein [Homo sapiens]		265020, 33657023, 264692, 264693, 264404,	_
1 AF13295 - (AF132954) CGI-20 P_VOLCA - SULFATED :IN 185 (SSG 185) 1 AF13295 - (AF132954) CGI-20	i		UNCLASSIFIED	264691 264556 264566	1
1 AF13295 - (AF132954) CGI-20 P_VOLCA - SULFATED :IN 185 (SSG 185) 1 AF13295 - (AF132954) CGI-20	100	Protein sim. GBank		52644045, 265007, 264632	1
P_VOLCA - SULFATED SIN 185 (SSG 185) 1 AF13295 - (AF132954) CGI-20	∞ .	0679 gb AAD27729.1 AF13295 - (AF132954) CGI-20			
P_VOLCA - SULFATED SIN 185 (SSG 185) 1 AF13295 - (AF132954) CGI-20	·= 1-	[Homo sapiens]			
PSUF_VOLCA - SULFALED ROTEIN 185 (SSG 185) Bank 7729.1 AF13295 - (AF132954) CGI-20 IS]	, ₹	Protein Sim. GBank		35696286, 66714117, 264508, 264509,	
3ank 7729.1 AF13295 - (AF132954) CGI-20	er ti	azujspjeznag/jssgip_volca - sulfated ACE GLYCODROTEIN 185 /SSG 185/		56182435, 264512, 18108351, 264688,	
Jank 7729.1 AF13295 - (AF132954) CGI-20 Is]	. 1			2564186 264486	
	₩ :	Protein sim. GBank		264686, 264488, 264687, 264489, 264768,	
	ο .	0679 gb AAD27729.1 AF13295 - (AF132954) CGI-20		264769, 264689, 21906769, 35696286,	
20281099, 16108564, 3569605, 264509, 264509, 264509, 264906, 18108, 264509, 264509, 264509, 264509, 264509, 264509, 264509, 264509, 26500, 26500, 26500, 26500, 26500, 26500, 26500, 26500, 26500, 264509, 26451, 264509, 26459, 26459, 26459, 26459, 26459, 26459, 1810836, 18108, 264564, 264448, 264565, 26478, 264565, 264789, 264565, 264789, 264565, 264789, 264565, 264789, 264565, 264789, 264565, 264789, 264565, 264789, 264568, 264488, 264568, 264488, 264568, 264488, 264568, 264488, 264568, 264488, 264568, 264488, 264568, 264488, 264568, 264488, 264588, 264588, 264	≡	l Homo sapiens]		35695917, 264259, 264691, 264692, 264693,	
264909, 264909, 19102 264907, 66712602, 264908, 264 18108374, 18108376, 3599642; 264510, 264511, 265006, 26500 265008, 264910, 264511, 264036, 264639, 264639, 264598, 264489, 264565, 264288, 264489, 264568, 264488, 264568, 264488, 264588, 264488, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588, 264488, 264588,				20281099, 18108364, 35696052, 264508, 264600, 264606, 264606, 48460378, 26460	
18108374, 18108376, 35696422 264510, 264511, 265006, 26500 265008, 264910, 264531, 26463, 264636, 264636, 264636, 264636, 264636, 264636, 264636, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 18108385, 18 265011, 26504, 26468, 264689, 18108385, 18 26504, 26468, 26468, 26468, 26468, 264688, 26468, 264688, 264681, 264688, 264681, 264688, 264588, 264481, 264686, 264288, 264481, 264588, 264588, 264481, 264588, 264588, 264588, 264588, 264588, 264588, 264588, 264588, 264588, 264588, 264588, 264584, 264588, 264584, 264588, 264584, 264588, 264584, 264588, 264588, 264584, 264588, 264584, 264584, 264588, 264584, 264588, 264588, 264588, 264584, 264588, 264584, 264588, 264584, 264588, 264584, 264588, 264588, 264588, 26				204303, 204303, 204300, 10100370, 20402(264907, 66712502, 264908, 264909.	
264510, 264511, 265006, 26500 265008, 264910, 264631, 26463 264635, 264591, 264636, 26463 264638, 26459, 264639, 26463 264738, 26459, 18108385, 18 265011, 26476, 26458, 26458 264764, 264586, 264288, 26448				18108374, 18108376, 35696423, 35695855,	
265008, 264910, 264631, 264631, 264636, 264910, 264636, 264910, 264636, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 18108395, 18708595, 18708595				264510, 264511, 265006, 265007, 264512,	
264635, 264591, 264636, 26463 264639, 264593, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264539, 264539, 264539, 264539, 264563, 18108385, 18108				265008, 264910, 264631, 264632, 264634,	
264638, 264593, 264539, 264539, 264539, 264539, 264539, 264539, 264539, 264539, 264539, 264539, 264539, 264539, 264563, 18108395, 18108395, 18108395, 18108395, 18108395, 264563, 264448, 264565, 26478				264635, 264591, 264636, 264637, 264592,	
264758, 26458, 18108385, 181 265011, 264760, 264563, 18108 264564, 26448, 264565, 26478 264764, 264586, 264288, 2644				264638, 264593, 264639, 264594, 83373044,	-
265011, 264760, 264363, 18100 264564, 264448, 264565, 26478 264764, 264586, 264288, 26448				264758, 264596, 18108385, 18108387,	
264764, 264586, 264288, 26448				265011, 264760, 264563, 18108351, 264762, 264664, 264664, 264469, 264664	
204/04, 204200, 204200, 20440				204304, 204440, 204303, 204703, 204003,	
90F79C			-	204/04, 204300, 204288, 204480, 20430/,	

87770203 (3607, 360	1804 87770203 (3607, 3608) Novel Protein sim. GBank gil3879914 emb CAA98538.1 - (Z74043) predicted using Genefinder; cDNA EST			52646365, 22278997, 22278999, 264905, 264908, 264909, 264910, 21906754, 264766,
	EMBL:C13850 comes from this gene; cDNA EST EMBL:C11575 comes from this gene; cDNA EST yk343f4.5 comes from this name (Caenorhandtits elegans)			21906765, 21906768, 35695917, 265020, 265022, 264691, 264637, 264639, 22279000, 264564, 264566
30375 (3609, 361	95330375 (3609, 3610) Novel Protein sim. GBank gil53544 ref NP_006461.1 pEBBP - estrogen-responsive B hoverotein			29331824, 29331825, 29331826, 29331827, 29331828, 87168559, 264288, 264687, 52644229, 35696423, 264636, 60432113
133762 (3611, 361	94133762 (3611, 3612) Novel Protein sim. GBank gil4589676 dbj BAA76857.1 - (AB023230) KIAA1013 protein [Homo sapiens]	15	struct	264094, 264105, 264908, 35696423, 265006, 265007, 265008, 264555, 264592, 265011, 265018, 264369
86943032 (3613, 3614)	4)			29331824, 264908, 264910, 33657023, 263978
642711 (3615, 361	87642711 (3615, 3616) Novel Protein sim. GBank gild884079 emb CAB43235.1 - (AL050008) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264488, 35696286, 66714117, 35696052, 66712502, 264592, 60433438, 52644296,
				265010, 264683, 264369, 264689, 55811957, 35695917, 33657109, 35695763, 55810764, 18108379, 35696423, 35695856, 56182323, 264563, 264564, 264487
3321468 (3617, 361	95321468 (3617, 3618) Novel Protein sim. GBank gi 1916927 (U87965) - putative G protein [Mus musculus]		UNCLASSIFIED	264594, 55811150, 264686, 29148629, 29148784, 264690, 264629, 18108374, 264556, 264557, 264558
3096316 (3619, 362	88096316 (3619, 3620) Novel Protein sim. GBank gij1352944jspjP47179jVJ9P_YEAST - HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DAL5 INTERGENIC REGION PRECURSOR	ב	UNCLASSIFIED	264488, 35696052, 264905, 264906, 264907, 264908, 264909, 264511, 265009, 264910, 264592, 264594, 33657402, 264757, 264596, 264758, 264596, 264758, 264696, 264764, 264689, 264764, 264769, 264289, 264764, 264768, 264689, 264764, 264269, 264768, 264689, 264289, 264689, 264260, 264260, 264260, 264260, 264260, 264260, 264260, 264260, 264260, 264260, 264260, 26
				264084, 264769, 264789, 264689, 265021, 264690, 266691, 264691, 264691, 264691, 264631, 264632, 264632, 264634, 264634, 264637, 264638, 264638, 264639, 264636, 264636, 264638, 264639, 264563
3086272 (3621, 36 <u>;</u>	88086272 (3621, 3622) Novel Protein sim. GBank gij2134984 pir 137275 - death-Conta associated protein kinase (EC 2.7.1) - human Ank r	Contains protein domain (PF00023) - kinase Ank repeat	nase	264488, 264259, 264508, 264509, 264905, 264906, 264906, 264907, 56182435, 264511, 264512, 264910, 264758, 265011, 264600, 264690, 264686, 264680, 2646335, 264689, 264631, 264633, 264
79245772 (3623, 3624)	54)			29331822, 29331824, 265019, 18108351, 21906769

z (3625, 3626 _,	1813 88090972 (3625, 3626) Novel Protein sim. GBank gi 5051636 gb AAD38326.1 AF07372 - (AF073727) EH	Contains protein domain (PF01417) - glucoamylase ENTH domain	glucoamylase	56182575, 264259, 29331824, 66714117, 29331824, 66714117,
	domain-binding mitotic phosphoprolein [Homo sapiens]			264906, 254907, 264908, 66712502, 264909, 265007, 264910, 264591, 264593, 55812038, 265011, 265018, 264760, 264682, 264764, 2646893, 264768, 2
1.				. 1800/00, 21906/08, 204691, 204693, 18108374, 35695855, 264634, 264635, 264637, 264639, 264559, 22279000, 22279002, 264566
	88178047 (3627, 3628) Novel Protein sim. GBank gij3643608 (AC005395) - hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	264488, 35696286, 22278998, 264092, 264084, 35696286, 22278998, 264092, 264094, 264299, 29331822, 29331824, 29331825, 29331825, 29331825, 29331826, 29331827, 28331828, 264105, 264107, 52644045, 56182436, 265009, 60432229, 60433356, 87168474, 87168559, 264389, 264288, 21906765, 28565917, 265021, 265022, 33657023, 33657029, 18108374, 35696423, 264838
☆	85296473 (3629, 3630) Novel Protein sim. GBank gi 11778B sp P26770 CYA4_RAT ADENYLATE CYCLASE, TYPE IV (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE)		siruct	22278999, 264508, 264509, 264907, 264908, 264910, 265011, 264760, 264766, 264636
ହ	83738845 (3631, 3632) Novel Protein sim. GBank gi 1176623 sp P41846 YO96_CAEEL - HYPOTHETICAL 93.9 KD PROTEIN T20B12.6 IN CHROMOSOME III		UNCLASSIFIED	18108394, 18108397, 264509, 264907, 264908, 264009, 265009, 264591, 265011, 265017, 264687, 264689, 265022, 264691, 18108362, 18108368, 18108370, 18108374, 18108379, 264635, 264557, 264564, 264567
2	84095268 (3633, 3634) Novel Protein sim. GBank gij3766377 emb CAA21429 - (AL031907) hypothetical protein [Schizosaccharomyces pombe]	Contains protein domain (PF00400) - histone WD domain, G-beta repeat	histone	264488, 22278997, 22278999, 60432049, 29331822, 29331824, 60432289, 52644045, 60170831, 265017, 265018, 265019, 18108351, 264822, 25844229, 21906765, 21906767, 21906768, 52644129, 23657023, 33657109, 27486262, 18108370, 18108374, 50170344, 5618737, 37, 27, 27, 27, 27, 27, 27, 27, 27, 27, 2
<u> </u>	85806775 (3635, 3636) Novel Protein sim. GBank gil3879121 emb CAA94370 - (Z70310) predicted using Genefinder; Similarity to Mouse ankyrin (PIR Acc. No. S37771); cDNA EST EMBL:101923 comes from this gene; cDNA EST EMBL:D32335 comes from this gene; cDNA EST EMBL:D32723 comes from this gene; cDNA EST EMBL:D32723 comes from this	Contains protein domain (PF00023) - Iranscriptfactor Ank repeat	transcriptfactor	35896286, 60433356, 264758, 264369, 264686, 21906769, 264693, 264632
ĝ	Novel Protein sim. GBank gi 5031865 ref NP_005771.1 pLHFP - lipoma HMGIC fusion partner		UNCLASSIFIED	22278995, 22278997, 22278999, 264259, 29331826, 264908, 265007, 265008, 265009, 60432229, 265017, 265018, 265019, 264448, 264288, 21906768, 21906769, 265020, 18108384, 22279000, 22279002, 264567

1820	1820 87769455 (3639, 3640)	3640)				264905, 264907, 264594
1821		3642)				264907, 264768, 263978
1822		3644) No	91221523 (3643, 3644) Novel Protein sim. GBank gil4884130 emb CAB43272.1			22278995, 56994075, 22278996, 22278997.
		· -				29331826, 35696052, 29331828, 264908,
						29331830, 60170831, 264591, 264593,
						60433356, 264596, 265017, 265019,
						18108351, 264763, 264683, 21906765,
						21906767, 21906768, 21906769, 35695917,
						265020, 265021, 33657023, 18108364,
						18108370, 35695855, 22279000, 22279002
1823	1823 85522330 (3645, 3646)	3646)			UNCLASSIFIED	264488, 264259, 264511, 264288, 264768,
						264693, 35696423, 264634, 18108385,
						264486
1824		3648) No	86612025 (3647, 3648) Novel Protein sim. GBank gi[477072]pir] A48018 - mucin 7	Contains protein domain (PF00047) - UNCLASSIFIED	_	264907, 264908, 264909, 264511, 264631,
		pre	precursor, salivary - human	Immunoglobulin domain		264634, 264635, 264637, 264638, 264639,
		-				264758, 264566
1825	87430125 (3649, 3	3650) No (AL	87430125 (3649, 3650) Novel Protein sim. GBank gij3036803jembjCAA18493j - (AL022373) hypothetical protein (Arabidoosis thaliana)		UNCLASSIFIED	60432049, 264910, 264487
1826		3652) No	91723612 (3651, 3652) Novel Protein sim. GBank		ATPase associated	ATPase associated 52644507, 52645156, 52646842, 22278994.
		9i ¥i6	gi 4680685 gb AAD27732.1 AF13295 - (AF132957) CGI-23		ı	22278996, 56994075, 264259, 60432049,
		pro	protein [Homo sapiens]			52645080, 35696052, 66712502, 52644045,
						265008, 265009, 60432229, 60433356,
						60433438, 52646317, 52644296, 265011,
						87168559, 264448, 264288, 264369, 264688.
						52644229, 264689, 21906765, 21906768,
						265020, 60170615, 52644150, 33657023,
						27486262, 27486264, 27486265, 35695763,
	·					35696423, 35695855, 83373044, 87168518,
						264404, 22279002
1827	1827 81647212 (3653, 3654)	3654)				264758

1828 95074017 (2855, 3859) Nover Protein am Claim? Contains protein domain (PFD0 13) - Unicogene 569498, 52548642, 5618275, 2227899, 54093,
Contains protein domain (PF00113) - Oncogene Enot-ase Enot-a
protein domain (PF00113) - oncogene
HED CO.
264468, 52646842, 56182575, 22278996, 35696286, 22278997, 22278999, 264091, 2640991, 2640991, 2640997, 2640997, 2640997, 2640997, 2640997, 2640997, 2640997, 2640997, 2640997, 2640997, 2640997, 2640997, 2640997, 2640997, 264508, 264004, 264508, 264004, 265008, 264007, 265009, 26172502, 52644045, 265009, 26170831, 60432229, 265009, 26170831, 60432229, 265009, 26170831, 60432229, 265009, 26170831, 264761, 264509, 265009, 26170831, 264761, 264761, 264762, 264448, 264764, 26468, 264768, 18108358, 264768, 264768, 264768, 264768, 264768, 264768, 264768, 264768, 264768, 264768, 264769, 2647166, 3659688, 264769, 264769, 2647169, 2647169, 2647169, 2647169, 2647169, 2647169, 2647169, 2778999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22378994, 22278999, 22378999, 22278999, 22378999, 22278999, 22378994, 22278999, 22378999, 22378999, 22378999, 22378997, 22278999, 22378999, 22378997, 22278999, 21906769, 2

1833	1833 95314184 (3665, 3666) Novel Protein sim. G	Novel Protein sim. GBank gi[5174413 ref NP_006026.1 pCDC4 - CDC42-binding	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	kinase	22278994, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 60432289.
		protein kinase beta (DMPK-like)			29331827, 35696052, 29146499, 264508, 264509, 264908, 264907, 66712502, 264008
					52644045, 264909, 264512, 265008, 264591,
					264593, 60433356, 21906754, 33657084,
					265011, 265017, 264604, 265018, 265019,
					264681, 18108351, 264683, 264288, 264685,
					264766, 264687, 21906765, 21906766,
					Z1905/67, Z1906/68, Z1906/69, Z9148629,
					265020, 265021, 264690, 264692, 33657023,
					65274620, 33657182, 27486264, 33657349,
					652/4/91, 264634, 264635, 264556, 264557, 1
					204030, Z04039, 10100303, 303Z0460,
	\rightarrow				264563
183 45	80562790 (3667, 3668)				264259, 264907, 264689, 22279000,
1835	04135718 (3660 3670)				22279002
200				UNCLASSIFIED	22278998, 29331822, 29331826, 87168474,
					264603, 21906768, 263976, 35695855,
1836	_	87748450 (2671 2672) Mayol Destain sim Comit			83373044
<u></u>		INOVER PROTEIN SITE, GDBITK	Contains protein domain (PF00153) - transport	transport	29331825, 264908, 265019, 264764, 264686,
4037		Silving process of the complete process of the complet	Mitochondrial carrier proteins		21906765, 264635
202		94234297 (3573, 3574) Novel Protein sim. GBank	Contains protein domain (PF00443) - ubiquitin	ubiquitin	22278995, 29146499, 265006, 265008,
		gilsss44uujspjuz45/4juBPE_DROME - UBIQUITIN	Ubiquitin carboxyl-terminal hydrolase		265009, 265010, 264683, 21906765,
		THIS FORTER OF STREET CONTROL SEE 64E (UBIQUITIN	family 2		29148627, 29148629, 265020, 265021,
		INICES ERASE 64E) (UBIQUI IN-SPECIFIC			265022, 65274620, 18108370, 18108374,
	,	PROCESSING PROTEASE 64E) (DEUBIQUITINATING ENZYME 64E)		_	264556, 18108385
1838		94324369 (3675, 3676) Novel Protein sim. GBank gil1362599 pir A56154 - Abl	Contains protein domain (PF00568) -		29331822, 265017, 264760, 265020,
		substrate ena (enabled) - fruit fly (Drosophila melanogaster) WH1 domain	WH1 domain		83373044
1839	87456508 (3677, 3678)	87456508 (3677, 3678) Novel Protein sim. GBank gi 2117310 emb CAB09116.1 -		UNCLASSIFIED	60433438, 264601, 21906765, 21906766,
	j	(295520) hypothetical protein (Schizosaccharomyces pombe)			265021, 33657109, 264556
1840	87391708 (3679, 3680)	87391708 (3679, 3680) Novel Protein sim. GBank gil127560lspIP23249IMV10 MOUSE - PROTEIN MOV:10		UNCLASSIFIED	264693
1841	85818445 (3681, 3682)	85818445 (3681, 3682) Novel Protein sim GBank			
: !		gil4572464[qb]AAD23834,1[AF12365 - (AF123653) FF71			561825/5, 29331824, 29331826, 60433356, 264764 264288 22657022 263067
		[Homo sapiens]			204 (04, 204200, 3303/023, 20390/. 18108370, 18108374, 264631, 264555.
					264556, 264639

65274572, 29331822, 29331824, 29331825, 66714117, 29331826, 29331827, 264907, 264909, 52644045, 56182435, 264510, 265006, 265007, 285009, 264910, 60433356, 264757, 60433438, 56812038, 265017, 18108354, 264686, 264769, 38557023, 264628, 581108364, 38557109, 18108368, 264528, 5810764, 56182323, 18108384, 264563, 264564	264488, 56182435, 264769, 29331826, 29331828, 264511, 265006, 265007, 264910, 264631, 264509, 264690, 264636, 264564, 264691, 60432229, 60432049, 264259, 264629, 33657023, 264486, 264909, 264567, 264595, 264766	264908, 265022, 33657023, 87168518, 22279002	264259, 29331824, 264907, 264908. 66712502, 284510, 265007, 265008. 55812038, 265018, 21906765, 52544150. 33657109, 264555, 264556, 264557, 25182323, 18108382, 83373044, 18108385, 264564	264905, 264908	22278997, 264259, 29331824, 264909, 18108351, 263974, 22279002	29331825, 29331826, 29331827, 265017, 264683, 264288, 264766, 264768, 21906767, 21906768, 264692, 22279002	56182575, 265018	60432049, 264908	29331824, 35696052, 264910, 60433438, 264688, 35695917, 265020, 52644150, 65274620, 52644332
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	transcriptfactor	UNCLASSIFIED	transcriptfactor	tm7	ATPase_associated	struct	dna_rna_bind
			Contains protein domain (PF00628) - transcriptfactor PHD-finger	Contains protein domain (PF00008) - UNCLASSIFIED EGF-like domain	Contains protein domain (PF00010) - transcriptfactor Helix-toop-helix DNA-binding domain		Contains protein domain (PF00008) - ATPase_associated 56182575, 265018 EGF-like domain	Contains protein domain (PF00169) - struct PH domain	Contains protein domain (PF00096) - dna_rna_bind Zinc finger, C2H2 type
1842 90992645 (3683, 3684) Novel Protein sim. GBank gil 1326268 (U58728) - C54H2.1 gene product [Caenorhabditis elegans]		8744764 (3687, 3688) Novel Protein sim. GBank gi[2496887[sp Q09232]YQ22_CAEEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III	Novel Protein sim. GBank gi[1175494 sp Q09619 YAC5_SCHPO - HYPOTHETICAL 45.0 KD PROTEIN C16C9.05 IN CHROMOSOME I	84287872 (3691, 3692) Novel Protein sim. GBank gil3881080 emb CAA21739 - (AL032657) similar to EGF-like domain; cDNA EST yk299a12.3 comes from this gene; cDNA EST EMBL:D35398 comes from this gene; cDNA EST yk331h6.5 comes from this gene; cDNA EST yk299a12.5 comes from this gene; cDNA EST yk467g8.5 c	B7821497 (3693, 3694) Novel Protein sim. GBank gij5059323[gb AAD38967.1 AF15152 - (AF151522) hairy and enhancer of split related-1 [Homo sapiens]	86789360 (3695, 3696) Novel Protein sim. GBank gij5701854 emb CAB52191.1 - (AJ245417) G5b protein [Homo sapiens]	84287874 (3697, 3698) Novel Protein sim. GBank gi4503665[ref]NP_001989.1 pFBLN - fibulin 2 precursor	86689650 (3699, 3700) Novel Protein sim. GBank giļ4589582 dbj BAA76813.1 - (AB023186) KIAA0969 protein [Homo sapiens]	95419789 (3701, 3702) Novel Protein sim. GBank gi[220637]db] BAA01477 - (D10627) zinc finger protein [Mus musculus]
90992645 (3683, 3684)	95292692 (3685, 3686)	87444764 (3687, 3688)	95096673 (3689, 3690)	84287872 (3691, 3692)	87821497 (3693, 3694)	86789360 (3695, 3696)			95419789 (3701, 3702)
1842	1843	1844	1845	1846	1847	1848	1849	1850	1851

(AJ006278) acetylglucosaminyltransferase-like protein [Mus 264910, 264758, 264693, 264637, 18108381, musculus] Musculus] 94324456 (3713, 3714) Novel Protein sim. GBank gil4322670[gblpAD16120] - (AF094568) dentin phosphoryn [Homo sapiens] CAF094588, 265020, 265022, 33657109, control of the control of
A l Pase_associated

1858	87628311 (3715, 3716)	1856 87628311 (3715, 3716) Novel Protein sim. GBank gil4981903 gb AAD36415.1 AE00178 - (AE001788) ribosomal protein S15 (Thermologa maritimal	Contains protein domain (PF00312) - ribosomalprot Ribosomal protein S15		264757
1859		84407464 (3717, 3718) Novel Protein sim. GBank gil4240317 dbj BAA74937.1 - (AB020721) KIAA0914 protein [Homo sapiens]			22278996, 29331824, 265007, 33109954, 265019, 264369, 21906768, 29148784, 27486261, 52644332, 22279002
1860	17929308 (3719, 3720)	17929308 (3719, 3720) Novel Protein sim. GBank gi 4009522 (AF099731) - Connexin 31.1 [Homo sapiens]			265019
1861	88086370 (3721, 3722)	88086370 (3721, 3722) Novel Protein sim. GBank gil 2143637 pir 184505 - calcium - Contains protein domain (PF00285)	Contains protein domain (PF00285) -		264887, 264259, 29331822, 29331824,
		dependent actin-binding protein - rat	Citrate synthase		29331825, 265007, 265009, 264591, 33109954, 265010, 265019, 264369, 264288
					264686, 264691, 264693, 27486264,
					18108370, 18108374, 263977, 55811576,
					56182323, 264639, 22279000, 22279002, 264482
1862	87372923 (3723, 3724)	87372923 (3723, 3724) Novel Protein sim. GBank	Contains protein domain (PF00069) - kinase	kinase	35696286, 264259, 87168474, 264369,
		gij125493 sp P07313 KMLC_RABIT - MYOSIN LIGHT CHAIN KINASE, SKELETAL MUSCLE (MLCK)	Eukaryotic protein kinase domain		21906766, 264558, 264563
1863		85775037 (3725, 3726) Novel Protein sim. GBank gij3820909 emb CAA09299 -		UNCLASSIFIED	264601, 264766, 29148627, 29148629,
		(AJ010642) Dof protein (Drosophila melanogaster)			264692, 264629, 264635
1864	85547832 (3727, 3728)	85547832 (3727, 3728) Novel Protein sim. GBank gil4322263 gb AAD15985 -	Contains protein domain (PF00754) - synthase	synthase	22278999, 264259, 264907, 265018,
		(AF077738) metallocarboxypeptidase CPX-1 [Mus	F5/8 type C domain		18108370, 264634, 264635, 264555, 264556,
		musculus)			264638, 18108387
1865	87740827 (3729, 3730)	Novel Protein sim. GBank			22278999, 264490, 29331822, 66714117,
		gilz495727[splQ93073]Y256_HUMAN - HYPOTHETICAL			66712502, 265006, 265007, 265008, 265009,
		PROTEIN KIAA0256			264591, 60433438, 265010, 265019, 264760, 264449, 264769, 30448627, 20448620
					265020, 265022, 18108385, 60432113
1866	_	87266816 (3731, 3732) Novel Protein sim. GBank gij5262617 emb CAB45748.11 -		kinase	18108374, 264769, 18108377, 21906765,
		(AL080157) hypothetical protein [Homo sapiens]			21906766, 35696423, 56182575, 21906769,
					29148629, 35696286, 35695917, 265021,
_					264510, 264511, 264512, 264534, 264535, 66476834, 62644460, 264666, 264604
					60433356, 29331822, 264559, 264595.
					29331824, 18108385, 21906754, 33657182,
					29331827, 35696052, 33656970, 87168518,
					265017, 60431602, 22279000, 264508,
	-				264509, 18108351, 264907, 264682, 264557, 18108372, 264765, 264486
1867	84579159 (3733, 3734)	84579159 (3733, 3734) Novel Protein sim. GBank gij3859930 (AF078096) -		UNCLASSIFIED	264094
		forkhead/winged helix-like transcription factor 7 [Homo saniens]			
	T		T		

264569, 264488, 35696286, 56994075, 264259, 29331822, 29331824, 29331825, 2569052, 29331822, 29331824, 29331825, 35696052, 29331826, 29146498, 224509, 264909, 264906, 264906, 264907, 264510, 264511, 264512, 265009, 264910, 33657402, 264584, 264766, 18108351, 264762, 264681, 264682, 264764, 264369, 264766, 264681, 264687, 264764, 264769, 264769, 264769, 264769, 264693, 23657109, 2748626, 264628, 18108370, 264639, 264638, 264638, 264639, 264635, 264638, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264638, 264638, 264659, 264565, 264686, 264567, 264569, 264486, 264567, 264569, 264486, 264567, 264569, 264486, 264567, 264569, 264486, 264567, 264569, 264567, 264569, 264567, 264566, 264486, 264567, 264569, 264567, 264569, 264486, 264567, 264569, 264567, 264567, 264569, 264486, 264567	264905, 264907, 264908, 265007, 264565, 264566	264766, 264691, 264692, 83373044	264509, 264905, 264908, 264909, 264910, 264762, 264687, 33657023, 264632	264908, 21906766, 18108370, 263974, 87168518	264908, 264910, 87168559, 21906766, 264636
ubiquitin	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED		- homeobox
Contains protein domain (PF01428) - ubiquitin AN1-like Zinc finger					Contains protein domain (PF00412) - homeobox LIM domain containing proteins
Novel Protein sim. GBank gi 4510345 gb AAD21434.1 - (AC006921) unknown protein [Arabidopsis thaliana]	95310883 (3757, 3758) Novel Protein sim. GBanik gjl4929643[gb]AAD34082.1 AF15184 - (AF151845) CGI-87 protein (Homo sapiens)	1880 91012978 (3759, 3760) Novel Protein sim. GBank gil1550765[emb]CAA69283] - (Y08026) immune associated protein 38 [Mus musculus]	80214949 (3761, 3762) Novel Protein sim. GBank gilg3144 pir B40505 - hypothetical protein - suid herpeswins 1 (strain Indiana- Funkhuser or Becker)	86582450 (3763, 3764) Novel Protein sim. GBank gil2384956 (AF022985) - No definition line found [Caenorhabditis elegans]	1883 94216817 (3765, 3766) Novel Protein sim. GBank gil1351218 sp P47226 TES2_MOUSE - TESTIN 2 (TES2) [CONTAINS: TESTIN 1 (TES1)]
1878 95351056 (3755, 3756) Novel Protein sim. (AC006921) unkno	95310883 (3757, 3758)	91012978 (3759, 3760)	80214949 (3761, 3762)	86582450 (3763, 3764)	94216817 (3765, 3766)
1878	1879	1880	1881	1882	1883

95310895 (3767, 3769) Novel Protein sim. GBank gigase43/gpt/AC3402 (1/45) CGI-87 protein (Homo sapiens) 96244280 (3769, 3770) Novel Protein sim. GBank (1/45) GBank gigase418/gpt/AC371 - VERPRCLIN 96674062 (3771, 3772) Novel Protein sim. GBank gigase418/gpt/AC4023711 - (AFCA52841) Novel Protein sim. GBank gigase418/gpt/AC4023711 - (AFCA52804 (3773, 3773) Novel Protein sim. GBank gigase418/gpt/AC402804 (3775, 3778) Novel Protein sim. GBank gigase418/gpt/AC402804 (3775, 3778) Novel Protein sim. GBank gigase418/gpt/AC403604 (3775, 3778) Novel Protein sim. GBank gigase418/gpt/AC403604 (3775, 3778) Novel Protein sim. GBank gigase418/gpt/AC40369 (3773, 3778) Novel Protein sim. GBank gigase418/gpt/AC40369 (3773, 3780) Novel Protein sim. GBank gigase418/gpt/AC4036 (3773, 3780) Novel	264488, 18108394, 56181686, 22278998, 60432049, 264259, 29331822, 29331824, 29331824, 29331822, 29331824, 29331825, 29331824, 29331825, 29331824, 29331825, 264508, 264508, 264909, 264909, 264909, 264909, 264909, 264909, 264909, 264910, 264511, 264500, 265017, 264500, 265011, 87168559, 264760, 265017, 264604, 265019, 264605, 264609, 264764, 264288, 264369, 264766, 264689, 264766, 264689, 264769, 264689, 264769, 264690, 264691, 33657023, 264628, 60431528, 264629, 263973, 18108374, 55810764, 55811576, 35696423, 3569832, 264639, 264638, 264638, 264638, 264639, 264638, 264639, 264565, 264565, 264565, 264566, 264488, 264567	56182575, 264259, 264905, 264909, 265008, 264596, 264766, 26502, 264628, 60431528, 264634, 56526486, 264080, 264563			263978	264906, 55812038, 264758, 265010, 265018, 265019, 18108351, 264288, 21906765, 21906769, 265021, 33657023, 33657109, 56182323, 83373044, 18108385, 22279000, 22279002	18108398, 29147620, 264907, 265009, 264600, 265018, 18108351, 264288, 264689, 21906765, 21906768, 21906769, 264691, 264692, 264693, 264628, 18108370, 264636, 264558, 264404
1pF15184 - (AF151845) CGI-87 RP1_YEAST - VERPROLIN cgi[2854158]gb AAC02577.1 - line found [Caenorhabditis 3.1 pCPNE - copine VI (neuronal) cgi[38319931 emb CAB10841 - anoma-Associated Antigen MAGE kgi[1083308 pir [A5659 - otein - mouse (fragment) k gi[4240195 db BAA74876.1 -		UNCLASSIFIED		ATPase_associated		nelicase	UNCLASSIFIED
9885 (3767, 3788) Novel Protein sim. GBank gil929643[gplAAD34082.14F15184 - (AF151845) CGI-87 protein [Homo sapiens] protein [Homo sapiens] gil2507155[splP37370]VRP1_YEAST - VERPROLIN gil2507155[splP37370]VRP1_YEAST - VERPROLIN gil2507155[splP37370]VRP1_YEAST - VERPROLIN (AF04564) Novel Protein sim. GBank gil2354158[gb]AAC02577.11 - (AF04564) Novel Protein sim. GBank gil3319931[emb]CAB108411 - (28804 (3775, 3776) Novel Protein sim. GBank gil3319931[emb]CAB108411 - (28804 (3775, 3776) Novel Protein sim. GBank gil3319931[emb]CAB108411 - (28804 (3775, 3776) Novel Protein sim. GBank gil33319931[emb]CAB108411 - (28804 (3775, 3776) Novel Protein sim. GBank gil408308[pir][A56559 - enhancer-trap-locus-1 protein - mouse (fragment) (48020660) KIAA0853 protein [Homo sapiens]				Contains protein domain (PF00168) - C2 domain	Contains protein domain (PF01454) - MAGE family	Contains protein domain (PF00176) - SNF2 and others N-terminal domain	
9885 (3767, 3768) 4280 (3769, 3770) 4082 (3771, 3772) 9139 (3773, 3774) 5783 (3777, 3778)	Novel Protein sim. GBank gij4929643 gbpAD34082.1 AF15184 - (AF151845) CGI-87 protein [Homo sapiens]	Novel Protein sim. GBank gil2507155[sp P37370]VRP1_YEAST - VERPROLIN	Novel Protein sim. GBank gi[2854158]gblpAc02577.1 - (AF045841) No definition line found [Caenorhabditis elegans]	Novel Protein sim, GBank gi 5174421[ref NP_006023.1 pCPNE - copine VI (neuronal)	Novel Protein sim. GBank gi[3319931 emb CAB10841 - (298046) dJ14O9.2 (Melanoma-Associated Antigen MAGE LIKE) [Homo sapiens]	Novel Protein sim. GBank gij 1083308 pir A56559 - enhancer-trap-locus-1 protein - mouse (fragment)	Novel Protein sim. GBank gi 4240195 db BAA74876.1 - (AB020660) KIAA0853 protein [Homo sapiens]
1885 8764 1886 8687 1886 8687 1889 9125 1890 8762	95310885 (3767, 3768)						

1891 87013895 (3781, 378 <u>2)</u>			UNCLASSIFIED	264686, 264768, 264687, 264693, 29331822, 29331824, 264508, 264905, 264908, 18108370, 264508, 264908, 264909, 18108379, 265007, 265008, 264910
				264632, 264591, 264639, 264596, 18108384, 265010, 265011, 264601, 264055, 264563, 264369
3, 3784)	87642825 (3783, 3784) Novel Protein sim. GBank gil5689535 dbj BAA83051.1 - (AB029022) KIAA1099 protein IHomo sapiensl	Contains protein domain (PF01412) - struct Putative GTP-ase activation protein	struct	22278995, 264509, 87168559, 18108351, 264448, 264682, 266020, 264603, 19108274
	-	for Art		22279000
88533826 (3785, 3786)			aminin	264569, 65274572, 22278997, 22278999,
				264259, 29331822, 29331824, 66714117,
_				29331826, 264906, 265006, 265008, 265009,
_				204592, 205018, 204681, 264448, 264683, 18108354 264360 264684 264686 264766
				264687, 264689, 21906768, 265020, 264709,
_				60170615, 52644150, 264690, 264691,
				264692, 33657023, 264693, 33657109,
				264628, 18108374, 35695855, 264630,
				264632, 264634, 264557, 264558, 60170394,
10000				18108381, 18108385, 22279000
00909120 (3787, 3788)				264508, 264905, 264906, 264907, 264594,
				264684, 264690, 264692, 264630, 264635,
10020				264636, 264639, 264563
9, 3790)	ord31081 (3708, 3790) Novel Protein sim. GBank gij5262574jemb CAB45729.1 -	Contains protein domain (PF00435) -		56182575, 264259, 60432289, 29331826,
	(ALVOUTS) inypointerical protein [Momo Sapiens]	Spectrin repeat		264107, 264905, 264908, 264910, 60170831,
				264758, 265010, 265018, 264448, 264288,
				264768, 33657109, 264628, 55810764,
10000				18108379, 264634, 56182323, 56526486
6561 3555 (3791, 3792)		Contains protein domain (PF00627) - UNCLASSIFIED UBA domain	UNCLASSIFIED	264907, 265008, 264682, 264686, 21906768, 264630, 264631
3, 3794)	80565569 (3793, 3794) Novel Protein sim. GBank		Cadharin	264250
	gi728836[sp P39193]ALU6_HUMAN - IIII ALU SUBFAMILY SP WARNING ENTRY IIII			60450
5, 3796) 	87617637 (3795, 3796) Novel Protein sim. GBank		helicase	22278996, 22278998, 22278999, 29331824,
	01-AOM NIETONA - ESOOMI-OLAMISETS Ideigos TILIS			29331825, 60432289, 29331827, 35696052,
				29331828, 265008, 265019, 264681, 264682, 264448, 264369, 62644220, 21006766
				201111, 201000, 32011223, 21900103, 31006760, 2100676
				55810764, 22279000
7, 3798)	86673097 (3797, 3798) Novel Protein sim. GBank gi[2909819 (AF031548) - lenthrocyte membrane divrocardain Bh60 (Llumo aminos)	Contains protein domain (PF00909) - glycoprotein	glycoprotein	264259, 264508, 264909, 60432229, 264769,
10086 00		Animoment I ansporter Family		21906/65, 21906/69
, 2000)	Nover Frotein Sim. Goank gift 102881 (AFU17250) - vitellogenin precursor (Oreochromis aureus)		UNCLASSIFIED	264683

264488, 52644507, 52645156, 18108396, 52646365, 52646842, 18108397, 56182575, 22278994, 22278994, 52937825, 56994075, 35696286, 22278994, 22278998, 264490, 60432049, 264229, 29331825, 56744107, 29331826, 60432289, 29331827, 29331826, 60432289, 29331827, 29331828, 264906, 264907, 29331830, 29146498, 2654006, 264907, 29331830, 264908, 264909, 264112, 265006, 264512, 265008, 264910, 265009, 60170831, 60432229, 60433356, 33657402, 60433438, 55812038, 264429, 21906764, 21906767, 265001, 8168859, 265017, 265010, 265011, 8168859, 265017, 265021, 264481, 265017, 265021, 264481, 264288, 264686, 264429, 21906765, 21906766, 21906765, 21906766, 21906767, 265021, 264441, 265021, 265031, 2644332, 2665031, 2644332, 2665031, 264636, 264381, 2659643, 3569565, 60431850, 264636, 264931, 26644332, 6043180, 264636, 264931, 26644332, 6043144, 18108385, 87168518, 604311394, 18108385, 87168518, 60431139	264107, 263976	264259, 29331825, 29331827, 264508, 264907, 265009, 60170831, 60433356, 60433438, 264759, 21908754, 264448, 264288, 265021, 265022, 33657023, 264693, 55811576, 264555, 264556, 22279000	264509, 264905, 264906, 264907, 264908, 265007, 264910, 264686, 264768, 264687, 264769, 264693, 264628, 18108374, 264634, 264636, 264637, 264565	264488, 264768, 264769, 264689, 29148629, 35695917, 35696286, 264259, 264692, 18108362, 33657023, 29331824, 33657109, 29146499, 264508, 264509, 264905, 264907, 66712502, 264908, 264909, 2664423, 35696423, 264510, 264511, 264512, 264910, 264634, 264635, 264637, 264638, 33657402, 264768, 85658542, 264602, 264760, 264761, 264482, 264563, 264764, 264566, 264288, 264764, 264566, 264288, 264764, 264566, 264288, 264764, 264566, 264288, 264766, 264766, 264286, 264288, 264766, 264566, 264288, 264766, 264766, 264566, 264288, 264766, 264566, 264288, 264766, 264766, 264288, 264766, 264766, 264483, 264766, 264760, 264766, 264760, 264766, 264760, 26476	264637
transport	dna_rna_bind	UNCLASSIFIED	struct		UNCLASSIFIED
Contains protein domain (PF00242) - transport DNA polymerase (viral) N-terminal domain				Contains protein domain (PF00293) - Bacterial mutT protein	
1901 95196647 (3801, 3802) Novel Protein sim. GBank giljs85959 sp P38378 S81A_RAT PROTEIN SEC61 ALPHA SUBUNIT	80202013 (3803, 3804) Novel Protein sim. GBank gil4426613 gb AAD20451 - (AF098796) SLM-1 [Mus musculus]	87778554 (3805, 3806) Novel Protein sim. GBank gi 3747107 (AF095741) - unknown [Rattus norvegicus]	80434213 (3807, 3808) Novel Protein sim. GBank gil1352911 sp P47147 YJ80_YEAST - HYPOTHETICAL 80.2 KD PROTEIN IN CPA2-NNF1 INTERGENIC REGION	95351140 (3809, 3810) Novel Protein sim. GBank gi 3043714 dbj BAA25521 - (AB011167) KIAA0595 protein [Homo sapiens]	
95196647 (3801, 3802)		87778554 (3805, 3806)	80434213 (3807, 3808)	95351140 (3809, 3810)	12763822 (3811, 3812)
1901	1902	1903	1904	1905	1906

65274572, 22278996, 35696286, 22278998, 22278999. 264259, 60432049, 29331822, 29331825, 60424269, 29331828, 35696052, 264593, 60433356, 21906754, 55811386, 85658542, 87168559, 265018, 264681, 264684, 264288, 21906765, 21906768, 265020, 265022, 264690, 52644150, 264692, 33657103, 33657109, 35695855, 264636, 264638, 264648, 264648, 264648, 264648, 264648, 264648, 264648, 264648, 264648, 264648, 264648, 264648, 26	2278995, 2227896, 22278997, 22278998, 22278995, 22278996, 22278997, 22278998, 22278997, 22278998, 22278997, 22278998, 22278997, 22278999, 22278999, 2643249, 264326, 29331822, 60432289, 29331827, 29331822, 2617502, 264906, 264907, 264908, 6182435, 264907, 264908, 6182435, 264907, 264908, 6182435, 264907, 264908, 264908, 264909, 6182435, 264691, 265011, 265017, 26404, 265018, 265019, 264805, 55811150, 264681, 26448, 26428, 264605, 55811150, 264681, 21906765, 21906766, 21906765, 21906766, 21906767, 265020, 265022, 264691, 264692, 18108364, 65274620, 33657109, 33657349, 3569423, 264632, 264656, 264856, 264868, 264868, 264868, 264869, 18108364, 264867, 264867, 264891, 264831, 264632, 264556, 264565, 264486	264259, 264508, 264906, 264907, 264907, 264908, 264511, 264910, 264593, 264758, 264764, 264766, 18108370, 264634, 264637, 264486	35696286, 22278996, 22278999, 35696052, 264509, 264900, 264500, 264901, 264001, 264512, 264908, 264758, 265011, 264601, 264602, 264604, 264768, 264761, 264764, 264788, 264768, 264687, 264628, 18108370, 264629, 18108372, 18108374, 35695855, 264631, 264636, 264636, 264636, 264636, 264636, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264656, 264566, 264486
hydrolase	UNCLASSIFIED	UNCLASSIFIED	- hотеорох
Contains protein domain (PF00561) - Inydrolase alpha/bela hydrolase fold			Contains protein domain (PF00412) - homeobox LIM domain containing proteins
Novel Prolein sim. GBank gil4929585[gb AAD34053.1 AF15181 - (AF151816) CGI-58 protein [Homo sapiens]	95313641 (3815, 3816) Novel Protein sim. GBank gil3986770 (AF109906) - NG22 [Mus musculus]	85514505 (3817, 3818) Novel Protein sim. GBank gij2224653 dbj BAA20813 - (AB002354) KIAA0356 [Homo sapiens]	94216821 (3819, 3820) Novel Protein sim. GBank gi 1351218 sp P47226 TES2_MOUSE - TESTIN 2 (TES2) [CONTAINS: TESTIN 1 (TES1)]
1907 95351144 (3813, 3814) Novel Protein sim. GBank gil4929585[gb AAD34053. protein [Homo sapiens]	1908 95313641 (3815, 3816) N	1909 85514505 (3817, 3818) N	1910 94216821 (3819, 3820) 6

1911	1911 91725345 (3821, 3822) Novel Protein sim. GBank	Novel Protein sim. GBank	Contains protein domain (PF01119) - nuclease	nuclease	18108394, 56182575, 56182181, 29331826,
		gildebassalge AAL30164.1 AC00633 - (AC006330) https://doi.or/.ndain/Lamo.goniana	DINA mismatch repair protein		29331827, 33535970, 254906, 253007.
		nypomencai protein (nomo sapiens)			204391, 33612036, 07100339, 204446, 264369, 21906765, 21906768, 265022
					264691, 264693, 18108365, 55811576,
					264556, 18108385, 18108388
1912	-	95413519 (3823, 3824) Novel Protein sim. GBank gij5689439 dbj BAA83003.1 -	Contains protein domain (PF00098) - UNCLASSIFIED	UNCLASSIFIED	18108397, 56182575, 56181686, 22278994,
		(AB028974) KIAA1051 protein [Homo sapiens]	Zinc finger, CCHC class		22278995, 56994075, 22278996, 22278999,
					264259, 29331822, 29331824, 56182181,
					29331825, 66714117, 35696052, 264905,
					264906, 264907, 264908, 52644045.
					56182435, 265007, 265008, 264910, 265009,
					264591, 264596, 65274444, 55811386,
_					87168474, 265011, 87168559, 265018,
					265019, 264760, 18108351, 264681, 264369,
					264684, 264288, 264686, 264768, 21906765,
					21906766, 21906767, 21906768, 21906769,
					55811957, 265020, 265021, 265022,
					60170615, 264692, 33657023, 264693,
					18108376, 55811576, 35696423, 65274791,
					264637, 56182323, 83373044, 56526486,
					22279002, 264563, 264566
1913	-	95305546 (3825, 3826) Novel Protein sim. GBank			56182575, 22278994, 22278995, 56994075,
		gil5032245 ref NP_005665.1 pZNF2 - zinc finger protein			22278996, 22278998, 22278999, 29331826,
		(C2H2) homotogous to mouse MOK-2			29331827, 265006, 55812038, 265010,
			٠		265017, 265018, 265019, 264681, 18108351,
					264683, 264764, 264369, 264288, 264685,
					264686, 264769, 21906765, 21906766,
					21906768, 21906769, 55811957, 265020,
					265022, 264691, 55811576, 264634, 264635,
					264638, 56182323, 83373044, 18108385
1914	_	83423982 (3827, 3828) Novel Protein sim. GBank gil4589604 dbj BAA76824.1 -	Contains protein domain (PF00036) - struct	struct	56182575, 29331824, 35696052, 264906,
		(AB023197) KIAA0980 protein [Homo sapiens]	EF hand		264908, 264828, 264909, 264592, 264758,
					87168559, 18108351, 18108354, 264684,
				_	264686, 33657023, 264693, 264628, 264631,
	_				264632, 264634, 264635, 264639
1915	95340459 (3829, 3830)	95340459 (3829, 3830) Novel Protein sim. GBank gi[5689415]dbj BAA82991.1 -		UNCLASSIFIED	264259, 29331824, 29331826, 29331827.
		(AB028962) KIAA1039 protein [Homo sapiens]			264508, 264909, 265009, 265017, 265019,
					264768, 264769, 264689, 264628, 264635,
	-				264637, 264639, 83373044, 264565
1916	79640761 (3831, 3832)	and the second s			264693, 264639

1	1 000 0000 000 0000			
	1917 87821580 (3833, 3834) Novel Protein sim. (AB028950) KIAA1	Novel Protein sim. GBank gijobbys99 (dbj BAAAb2979.1 - (AB028950) KIAA1027 protein [Homo sapiens]	struct	264799, 264689, 21906765, 21906768, 22278996, 264259, 264691, 264693, 29331824, 29331825, 29331826, 29331828,
				264905, 264510, 264638, 264910, 264908, 264909, 264510, 264630, 264910, 264634,
				264635, 264636, 264637, 264638, 263981, 264639, 264758, 18108385, 21906754,
				265011, 264604, 264563, 18108351, 264762,
	10000 10000			264763, 264565, 264764, 264766
8181	95302795 (3835, 3836) Novel Protein sim.	Novel Protein sim. GBank	struct	264488, 18108392, 18108357, 21906765,
		[Ratter solvedicus]		22278994, 35696286, 35695917, 22278996.
				22278997, 265021, 265022, 264534, 264690,
				264691, 264692, 33657023, 264693,
				29331824, 29331825, 33657109, 29331826,
				52645129, 35696052, 29331828, 27486262,
_				[Z/486Z64, 35695/63, Z64508, Z64905,
				264509,264906,264628,264907,18108370, 264908 264620 264000 18108372
				201900, 201029, 201909, 101009/2, 18108374, 263978, 35696423, 35695855
			-	264510, 264511, 265006, 265007, 264512.
				265008, 264631, 265009, 264910, 264634,
				264635, 264555, 264636, 264556, 264637,
				264557, 264593, 264638, 264594, 60170394,
				264595, 264559, 264596, 83373044, 264758,
				52646317, 18108385, 52644296, 56526486,
				07 1003 0, 2030 10, 2030 11, 07 100339; 364600 364604 364603 366047 364603
				264604, 265018, 264605, 263017, 264605,
				264482, 264564, 18108351, 264762, 264682,
				264565, 264448, 264764, 264566, 264486,
				264567, 264369, 264288, 264766, 264487,
1919	94143847 (3837, 3838) Novel Protein sim.	Novel Protein sim. GBank gil3878584lemblCAB01237I -	oxidase	22278997, 29331822, 265007, 60170831.
				60432229, 60433438, 264448, 264682,
		cDNA EST EMBL:C09753 comes from this gene		264288, 55811957, 33657023, 33657109,
		[Caenorhabditis elegans]		65274791, 56182323, 22279002
1920	91229953 (3839, 3840) Novel Protein sim.	Novel Protein sim. GBank gij1809231 (AC000115) - coded	UNCLASSIFIED	264510, 264511, 264512, 264566
		1or by human cDNAs R76043 (NID:g850725), R65857 (NID:a838405) and H12888 (NID:a877888) (Homo conjunct)		
5	79555226 (3841 3842)	4024 7055.6208 (3841 3842) Novel Priving George (im CBant CB	LINCI ACCIEIED	SEARON
;	י מפחקדים (חברי הי מיבי	oil4580997lablAAD24571.11AF12108 - (AF121081) cAMP	ONO DIVIDIO	Sector
		inducible 2 protein (Mus musculus)		

1930	1930 87889128 (3859, 3860) Novel Protein sim. GBank	Novel Protein sim. GBank		phosphatase	35696286, 29331828, 264905, 264907,
		gij1709230jsp P52963JNBL4_MOUSE - NBL4 PROTEIN			264908, 264909, 264511, 264910, 264758,
					264601, 265017, 265019, 264605, 264760,
					264764, 264766, 264686, 264769, 265022.
	┰				35696423, 264638, 60432113
1931		87797279 (3861, 3862) Novel Protein sim. GBank gi 404634 (U01840) -	Contains protein domain (PF00069) - kinase	kinase	264906, 264908, 60432229, 264758, 264764,
	┱	serine/threonine kinase [Mus musculus]	Eukaryotic protein kinase domain		264288, 265020, 264692, 264634, 264637
1932	\neg				264684, 264691, 264635
1933	11613668 (3865, 3866)			UNCLASSIFIED	284595
1934	84426360 (3867, 3868)			struct	56182575, 56182435, 264510, 264757,
		(AB022023) nonmuscle myosin heavy chain B [Bos taurus]			264758, 55812038, 55811386, 265018,
					55811150, 21906765, 264691, 264631,
١					264635, 264637
1835	87752511 (3869, 3870)			UNCLASSIFIED	264686, 265011, 264511, 264905, 18108351,
					264564, 264681, 264259, 18108370, 264566, 1 264764, 264369, 264595, 18108370, 264566, 1
1936		95414338 (3871, 3872) Novel Protein sim. GBank			60432289 265007 265010 265011 265019
_	-	gil4827040frefiNP 005110,1lpTRAP - thyroid hormone			33657109 18108374
		receptor-associated protein, 150 kDa subunit			
1937		94847141 (3873, 3874) Novel Protein sim. GBank gi[543187 pir S37771 - ankyrin,	Contains protein domain (PF00023) - kinase	kinase	85658542, 21906767, 35695917, 60170615,
į	_	erythrocyte - mouse	Ank repeat		264693, 33657109
1938	87403277 (3875, 3876)	87403277 (3875, 3876) Novel Protein sim. GBank	Contains protein domain (PF00888) - collagen	collagen	264488, 29146498, 264905, 264559
		gil4544431[gb]AAD22340.1[AC00695 - (AC006955) hypothetical protein (Arabidoneis thaliana)	Cullin family		
4000	04004079 (2977 2970)	Manual Destriction of the Proposition of the Propos		40.0000	
8581	91004978 (3877, 3878)	91004976 (3677, 3576) Novel Protein sim. GBank gij500858[dbj]BAA03210] - [(D14168) 50kDa lectin [Bombyx mori]		UNCLASSIFIED	65274572, 22278999, 60432289, 56182435, 60433356, 265017, 21906765, 21906766,
					21906768, 55811957, 27486264, 35696423.
3	_				00432113, 204304
		8/348810 (36/9, 3880) Nover Protein sim. Glank gij1946300 emb CAA/3132 - (Y12529) hypothetical protein [Silene latifolia]	Contains protein domain (PF00560) - struct Leucine Rich Repeat	struct	264488, 29331822, 264448, 264683, 264288, 265020, 33657023, 264631
1941	94147177 (3881, 3882)	94147177 (3881, 3882) Novel Protein sim. GBank gil4206386 (AF060570) - rig-1		UNCLASSIFIED	56994075, 22278998, 264259, 29331824,
		protein [Mus musculus]			29331827, 264905, 265008, 33657084,
					265017, 265018, 264288, 264687, 21906765,
					21906766, 21906767, 265020, 52644150,
					27486264, 83373044, 18108387, 60432113,
	_				22279002, 264565
1942	87641870 (3883, 3884)	87641870 (3883, 3884) Novel Protein sim. GBank		UNCLASSIFIED	264488, 18108398, 29331825, 27486261,
-		gil492/204[gb AAD33049.1[AF13391 - (AF133911) ARL-6 interacting protein-4 fMus musculus]			264509, 18108370, 18108374, 264482
1943	94325298 (3885, 3886)	94325298 (3885, 3886) Novel Protein sim. GBank	Contains protein domain (PF00400) - kinase	kinase	22278998, 29331822, 29331827, 35696052,
		gij3122952 sp 015736 TIPD_DICDI - TIPD PROTEIN	WD domain, G-beta repeat		264511, 265009, 264592, 60432229, 265017,
					265018, 265019, 264684, 264692, 33657109,
					65274791, 264636

UNCLASSIFIED 65274572, 22278994, 22278995, 22278996, 22278998, 22278999, 2264259, 52645080, 29331824, 29331826, 29331827, 66712602, 56182435, 264512, 264910, 60170831, 60432229, 60433356, 33657402, 6043348, 264596, 33109954, 21906754, 87168474, 87168559, 265017, 265018, 265019, 18108351, 264369, 264686, 264768, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906767, 21906767, 21906767, 21906767, 21906768, 3569591, 265021, 60170815, 264692, 33657023, 18108374, 35696423, 3565965, 264634, 60170394, 264639, 83373044, 18108385, 56526486, 87168518, 6047313, 27270000, 264669	UNCLASSIFIED 264488, 22278996, 264510, 264511, 18108351, 264488, 264486, 264567	UNCLASSIFIED 60432289, 29331827, 35696052, 265007, 265008, 60433356, 60433438, 264369, 56181562, 21906767, 52844150, 264693, 27486264, 264637, 87168518, 264563	mapolymerase 22278998, 264905, 264908, 264909, 264909, 264512, 264758, 264762, 264682, 264683, 264768, 264768, 264788, 264768, 264688, 21906768, 264693, 18108374, 3569585, 264635, 264637, 264637, 264637, 264637, 264637, 264637, 264637, 264637, 264637, 264637, 264637, 264637, 264637, 264637, 264637, 264638, 27279007, 264637, 264638, 264637, 264638, 264637, 264638, 264688,		264369 264488, 264688, 264689, 264259, 264259, 29331822, 33657109, 29331828, 264103, 264509, 18108370, 35695855, 264510, 265008, 265009, 33657402, 18108385, 265018, 264503, 55811150, 18108351, 264369, 564288, 18108351,		
Š	חאל	חאכ	leu	struct	Contains protein domain (PF00857) - UNCLASSIFIED Isochorismatase family	Contains protein domain (PF00916) - transport Sulfate transporter family	Contains protein domain (PF00106) - reductase short chain dehydrogenase
1944 94232958 (3887, 3888) Novel Protein sim. GBank gij 1799570 db BAA13432 - (D87671) TIP120 [Rattus norvegicus]	87641872 (3889, 3890) Novel Protein sim. GBank gil4927204 gb AAD33049.1 AF13391 - (AF133911) ARL-6 interacting protein-4 [Mus musculus]	8/443990 (3891, 3892) Novel Protein sim. GBank gi 2498104 sp Q27969 AD50_BOVIN - ADRENAL MEDULLA 50 KD PROTEIN	86438862 (3893, 3894) Novel Protein sim. GBank gil3914801 sp O54888 RPA2_RAT - DNA-DIRECTED RNA POLYMERASE I 135 KD POLYPEPTIDE (RNA POLYMERASE I SUBUNIT 2) (RPA135) (RNA POLYMERASE I 127 KD SUBUNIT)	95199174 (3695, 3896) Novel Protein sim. GBank gij5420387jemb CAB46679.1	87788531 (3899, 3900) Novel Protein sim. GBank gil3876766 emb CAA93466.11- (269637) predicted using Genefinder; Similarity to E.coli hypothetical protein YCAC (SW:YCAC_ECOLI) [Caenorhabditis elegans]	86988253 (3901, 3902) Novel Protein sim. GBank gij2626763 dbj BAA23424 - (AB008782) sulfate transporter [Arabidopsis thaliana]	87069775 (3903, 3904) Novel Protein sim. GBank gil4929633[gb AAD34077.1 AF15184 - (AF151840) CGI-82 protein [Homo sapiens]
м 94232958 (3887, 3888)		8/443990 (3891, 3892)	86438962 (3893, 3894)	1948 95199174 (3695, 3896) 1940 7640120 (3807 3808)	T		1952 87069775 (3903, 3904)

264259, 264558	264488, 263994, 18108394, 35696286, 22278998, 29331822, 66714117, 29331826, 29331822, 66714117, 29331826, 29331827, 264508, 264905, 264906, 264906, 264907, 264908, 264909, 264500, 265009, 265009, 265009, 265009, 265009, 265009, 265009, 265019, 265019, 265019, 264691, 35695917, 265020, 265021, 264691, 33657023, 264692, 33657109, 264628, 18108374, 35696423, 55811576, 35695855, 264630, 264639, 264556, 264639, 264559, 264639, 264569, 264639, 264639, 264569, 264639, 264569, 264639, 264569, 264659, 264659, 264659, 264659, 264659, 264659, 264659, 264569, 264659, 264569, 264659, 264569, 264659, 264569, 264659, 264569, 264659, 264569, 264659, 264569, 264659, 264569, 264659, 264569, 264659, 264569, 264659, 264659, 264569, 264659, 264669, 264659, 264659, 264659, 264659, 264659, 264659, 264659, 2646	264269, 2931822, 56181181, 29331824, 264269, 29331825, 56181686, 22278994, 22278999, 56182575, 56181686, 22278994, 22278999, 264269, 29331825, 29331827, 36596052, 264508, 29331830, 265008, 265009, 264691, 55812038, 87168474, 265017, 265018, 265019, 264448, 264766, 21906765, 21906765, 21906765, 21906765, 21906765, 3657109, 27486264, 33657349, 35695763, 60431850, 264635, 3659865, 264635, 60431850, 264630, 39337874, 364630, 36431850, 364630, 364630, 36431850, 364630, 364630, 364635, 36431850, 364630, 364630, 364635, 36431850, 364630, 364630, 364635, 36431850, 364630, 364635, 364635, 36431850, 364630, 364630, 364635, 364635, 36431850, 364630, 364635, 364635, 36431850, 364630, 364630, 364635, 364635, 36431850, 364630, 364635, 364635, 36431850, 364630, 364635, 364635, 36431850, 364630, 364635, 364635, 364635, 36431850, 364630, 364635, 36463	264488, 56182575, 22278996, 22278999, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 264506, 264906, 264908, 264909, 264511, 265007, 264910, 264591, 55812038, 265010, 265018, 18108351, 264768, 56181562, 264689, 21906768, 21906769, 265022, 264689, 264628, 18108374, 55810764, 55811576, 264639, 264639, 264633, 264633, 264633, 264633, 264633, 264633, 264633, 264633, 264635, 264635, 264635, 264637, 264639, 22279000, 222790002, 264563, 264564
complement	ubiquitin	UNCLASSIFIED	cadherin
Contains protein domain (PF00386) - complement C1q domain	Contains protein domain (PF00443) - ubiquitin Ubiquitin carboxyl-terminal hydrolase family 2		Contains protein domain (PF00028) - cadherin Cadherin domain
Novel Protein sim. GBank gi[1168715]sp P31721 C1QB_RAT - COMPLEMENT C1Q SUBCOMPONENT, B CHAIN PRECURSOR	Novel Protein sim. GBank gil4240271(dbj BAA74914.1 - (AB020698) KIAA0891 protein [Homo sapiens]		95092121 (3911, 3012) Novel Protein sim. GBank gi 1665821 dbj BAA13407 - (D87469) Similar to D.melanogaster cadherin-related tumor suppressor [Homo sapiens]
1953 20470371 (3905, 3906) Novel Protein sim. 9 9i 1168715 sp P317 SUBCOMPONENT		95308310 (3909, 3910)	95092121 (3911, 3912)
1953	4500 4500		1956

1957	1957 94326510 (3913, 3914) Novel Protein sim GRa	Novel Protein sim GBank ail4589674/4hilBAA78968 11			
		(AB023229) KIAA1012 protein [Homo sapiens]		ONCLASSIFIED	52646842, 56182575, 22278997, 22278998, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 29146898, 264593, 33657402, 33109954, 87188474, 265018, 294448, 264369, 264268, 21906769, 265020, 265021, 264692, 65527486264, 33657349, 27486265, 24566868
1958	95313902 (3915, 3916)	95313902 (3915, 3916) Novel Protein sim. GBank gil4240227 dbj BAA74892.1 - (AB020676) KIAA0869 protein [Homo sapiens]		UNCLASSIFIED	20278999, 264092, 264094, 264259, 26278999, 264092, 264094, 264259, 260432049, 264410, 264107, 264109, 264909, 264511, 60170831, 60432229, 21906754, 265010, 21906769, 35695917, 265920, 253967, 35695917, 35695423, 264634, 26557650, 263967, 265076, 356954520, 263967, 265076, 3569545, 3569545, 3569545, 3569545, 3569545, 3569545, 3569545, 3569545, 3569545, 3569545, 356955, 35
1959	85701470 (3917, 3918)	B5701470 (3917, 3918) Novel Protein sim. GBank gi[2281983 emb CAB10860 - (298056) hyypothetical protein [Schizosaccharomyces pombe]		ubiquitin	264593, 265019
96	80308608 (3919, 3920)	1950 80308508 (3919, 3920) Novel Protein sim. GBank gi[2274851[db] BAA21515 - (D64159) 3-7 gene product [Homo sapiens]		struct	264905, 264906, 264907, 264909, 265006, 265007, 264910, 264695, 265017, 264604, 265018, 18108351, 264764, 264369, 264768, 264768, 21906765, 18108368, 264639, 18108379, 264635, 264636, 264637,
\neg	16292607 (3921, 3922)				204030, 204480 264636
	91008385 (3923, 3924)			UNCLASSIFIED	65274572 264502 264503 255010 264503
	90936017 (3925, 3926)	90936017 (3925, 3926) Novel Protein sim. GBank gij3721653 dbj BAA33581 - (AB012933) acyl-CoA synthetase 5 [Rattus norvegicus]		eph	65274572, 18108398, 35696286, 29331825, 6043289, 29331827, 264828, 29331825, 265006, 265006, 265002, 33657023, 33657109, 2488585, 265685
	94317605 (3927, 3928)	94317605 (3927, 3928) Novel Protein sim. GBank gij5262638 emb CAB45757.1 - (AL080169) hypothetical protein [Homo sapiens]		cadherin	264488, 264092, 264259, 264509, 264905, 264488, 264092, 264259, 264509, 264905, 264509, 264909, 264906, 264511, 265007, 264908, 264909, 264510, 264511, 265007, 264009, 264910, 264592, 264593, 264604, 264604, 264605, 264766, 264766, 264768, 264768, 264691, 264691, 264693, 18108370, 264628, 264639, 264638, 264658, 264566, 26468, 264566, 264566, 264566, 264566, 264566, 264566, 264566, 264566, 264566, 264566, 264566, 264566, 264566, 264566, 264
1965	94317445 (3929, 3930) (Novel Protein sim. GBank gi 4107017 dbj BAA36294 - (AB001773) PEM-6 [Ciona savignyi]	Contains protein domain (PF01428) - ubiquitin AN1-like Zinc finger	ubiquitin	264488, 264567 264488, 264510, 264760, 264768, 264486

1000	04400010 10004 00001	AAAAA AAAAA AAAAA AAAAA AAAAAAAAAAAAAA			
3	710400 (3001, 3004)	gil4929707[qb]AAD34114.1IAF15187 - (AF151877) CGI-119 Uncharacterized protein family	Uncharacterized protein family	giycoprotein	222/0399, 204092, 204239, 2933/020, 2933/828, 29146498, 264595, 265011,
		protein [Homo sapiens]	-		264448, 18108354, 264288, 264684, 264766,
					264685, 264686, 265022, 264691, 264692,
					18108370, 18108377, 264555, 18108381,
					18108385, 264486, 264567
1967	87396123 (3933, 3934)	87396123 (3933, 3934) Novel Protein sim. GBank gil2957270 (AF044576) -	Contains protein domain (PF00388) - esterase	esterase	29331824, 265010, 265017, 264288,
		phospholipase C PLC210 [Caenorhabditis elegans]	Phosphatidylinositol-specific phospholipase C, X domain		21906764, 263981, 56526486
1968		88095641 (3935, 3936) Novel Protein sim. GBank gi[2564953 (AF030001) -	Contains protein domain (PF00008) - oncogene	oncogene	35696286, 264905, 264509, 264906, 264907,
		unknown [Mus musculus]	EGF-like domain		264908, 264909, 264511, 264512, 265008,
					264910, 265009, 264594, 264757, 264758.
					264604, 264605, 264760, 264762, 264682,
					264764, 264685, 264766, 264767, 264689,
					264691, 264693, 264628, 264629, 35696423,
					35695855, 264631, 264632, 264634, 264635,
					264636, 264637, 18108380, 264564, 264565,
					264566, 264567
1969	84328529 (3937, 3938)	84328529 (3937, 3938) Novel Protein sim. GBank gil2911274 (U20329) - spidroin 1		UNCLASSIFIED	22278995, 22278996, 35696052, 264906,
		[Nephila clavipes]			264908, 18108351, 264482
1970	80596049 (3939, 3940)	80596049 (3939, 3940) Novel Protein sim. GBank gil4050087 (AF109907) - S164		UNCLASSIFIED	264908, 264288, 264766, 264636
		[Homo sapiens]			
1971	94843914 (3941, 3942)	94843914 (3941, 3942) Novel Protein sim. GBank		collagen	264488, 264489, 22278998, 264259,
		gi[134206 sp P09593 SANT_PLAFV - S-ANTIGEN			60432049, 66714117, 29331826, 60432289,
		PROTEIN PRECURSOR			29331827, 35696052, 264508, 264905,
					264509, 264906, 264907, 264908, 264909,
					264510, 264511, 264512, 264910, 264591,
					264592, 60432229, 60433356, 264595,
					264596, 264600, 264604, 264605, 264760.
					18108351, 264448, 264764, 264288, 264766.
					264768, 264769, 21906765, 33657023,
					264692, 18108370, 264629, 35696423,
					65274791, 35695855, 264632, 264635,
					264555, 264636, 264637, 264638, 264639,
					18108385, 60432113, 22279000, 264563,
į					264564, 264565, 264566, 264486
7/61		8/645444 (3943, 3944) Novel Protein Sim. GBank gil4519623 dbjjBAA75671.1 -	Contains protein domain (PF01462) -		22278999, 264259, 29331822, 56182181,
		(AB017616) homologous to the yeast YGR163 gene [Mus	Leucine rich repeat N-terminal		60432289, 29331827, 52644045, 264909,
		musculus]	domain		265006, 264511, 265008, 52644296, 265018,
					265019, 264761, 264689, 21906768,
					21906769, 264691, 264693, 33657109,
					33657182, 264556, 52644332, 264558,
	_				60432113
1973	86395533 (3945, 3946)			UNCLASSIFIED	29331826, 264692, 35696423, 264631,
					264555, 264556, 264557, 264558, 264559
1974	80396629 (3947, 3948)	80396629 (3947, 3948) Novel Protein sim. GBank gij3309543 (AF036382) - MLL [Film: rubrines]		UNCLASSIFIED	264682, 264764, 264563
		foodings and and			

264488, 66714117, 29331826, 29331828, 56182435, 265006, 264757, 55812038, 265010, 265017, 264369, 55811957, 65274791, 35695855, 56182323, 60432113	264259, 35696052, 265018, 265020, 265021, 33657109, 56526486	264908, 264596, 265021, 264566	60170831, 264566	29331826, 29146498, 264905, 264907, 265007, 265009, 265010, 265018, 264686, 18108359, 21906768, 35695917, 265020, 60170615, 264693, 18108369, 18108370, 264631, 264635, 264556, 264558, 18108384, 22730100, 264556, 264558, 18108384,	264489	22278995, 22278997, 22278999, 264259, 29331825, 29331828, 29146498, 29146499, 264107, 264906, 264910, 264595, 21906754, 265010, 265017, 265018, 265019, 26448, 26428, 21906767, 33637023, 27486264, 18108370, 18108376, 264630, 264631, 264635, 18108385, 87168518, 27370010, 264832, 264636	65274572, 29331825, 35696052, 33656970, 264969, 265008, 55311386, 264760, 264686, 264661	22778996, 35696286, 264259, 29331822, 20281099, 29331824, 60432289, 29331827, 264509, 264905, 264907, 66712502, 264908, 5264905, 264907, 66712502, 264910, 265009, 264501, 264502, 264910, 265009, 264501, 264600, 264600, 264604, 264768, 264768, 264768, 264769, 264603, 264604, 264628, 264769, 264609, 264632, 3669517, 264690, 264633, 366955, 20281071, 264632, 264634, 264638, 264638, 264638, 264639, 264638, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 264639, 26279000, 22279000
UNCLASSIFIED	UNCLASSIFIED	потеорох	transcriptfactor	UNCLASSIFIED	ubiquitin	UNCLASSIFIED		transcriptfactor
								Contains protein domain (PF00096) - Iranscriptfactor Zinc finger, C2H2 type
			87447645 (3955, 3956) Novel Protein sim. GBank gi 103421 pir A33471 - transcription factor NTF1 - fruit fly (Drosophila melanogaster) (fragment)	Novel Protein sim, GBank gi[2244815 emb CAB10238.1 - (297336) hypothetical protein [Arabidopsis thaliana]	Novel Protein sim. GBank gil4759290 ref NP_004642.1 pUSP1 - Ubiquitin carboxyl- ferminal hydrolase, X-linked	B7606974 (3961, 3962) Novel Protein sim. GBank gil4406693 gb AAD20060 - (AF131849) Unknown [Homo sapiens]	90995367 (3963, 3964) Novel Protein sim. GBank gi 5689523 db BAA83045.1 - (AB029016) KIAA1093 protein [Homo sapiens]	Novel Protein sim. GBank gij3417297 (AC002310) - Unknown gene product [Homo sapiens]
1975 94316479 (3949, 3950)	95358914 (3951, 3952)	94632654 (3953, 3954) Novel Protein sim. gi[2499526]sp[Q07 COTRANSPORTE	87447645 (3955, 3956) Novel Protein sim. transcription factor melanogaster) (fra	87627709 (3957, 3959)	86577059 (3959, 3960) Novel Protein sim. gil4759290[ref]NP_ ferminal hydrolase.	87606974 (3961, 3962)	90995367 (3963, 3964)	95098668 (3965, 3966) Novel Protein sim. (Unknown gene proc
1975	9/61							1983

1984	85760989 (3967, 3968)	1984 (85760989 (3967, 3968) Novel Protein sim. GBank gi[2896695 emb CAA17174.1 -	<u>v</u>	synthase	264688, 21906766, 55811957, 56994075,
		(ALUZ1087) IddD14 [Mycobacterium iubercurosis]			200020, 200021, 22270999, 200022, 204209,
					29331822, 3365/182, 29146499, 264628,
					18108370, 264908, 264629, 55811576,
					35695855, 265006, 265007, 264591,
					21906754, 33657084, 265010, 265017,
					265019, 264288
1985		85636897 (3969, 3970) Novel Protein sim. GBank		glycoprotein	264760, 264288, 263978, 55811576, 264637,
		gi 5712131 gb AAD47379.1 AF12049 - (AF120499) DEM1 protein [Homo sapiens]		• •	56182323, 18108385, 264564
1986	80200507 (3971, 3972)			UNCLASSIFIED	264488, 264629
1987	$\overline{}$	Novel Protein sim. GBank	Contains protein domain (PF00069) -		22278999, 29331830, 265007, 265018,
		1 AF14457 - (AF144573) Mx-	Eukarvotic protein kinase domain		21906768, 33657023, 264692, 264693,
		interacting protein kinase PKM [Mesocricetus auratus]			18108377, 264635, 60170394, 22279002
1988	94122108 (3975, 3976)	-		UNCLASSIFIED	264905, 264906, 264907, 264908, 264909,
					264910, 264591, 264593, 264758, 264764,
					264686, 264768, 265021, 264692, 264628,
					264629, 35695855, 264630, 264635, 264636,
					264637, 264638, 264639, 264483
1989	91225225 (3977, 3978)		<u> </u>	tubulin	60432049, 60432289, 52644045, 56182435,
		pole body protein spc97 homolog GCP2 [Homo sapiens]			264112, 265007, 33657402, 52644229,
					21906765, 21906768, 21906769, 55811957,
					33657023, 263967, 33657109, 18108370,
					22279000, 22279002
1990		85699888 (3979, 3980) Novel Protein sim. GBank gi 5701727 dbj BAA83074.1 -			264508, 264757, 264764, 18108381
		(AB024729) alpha-1,3-D-mannoside beta-1,4-N-			
		acetylolucosaminyltransferase IV-homologue IHomo	•		
		sapiens			
1991	95353114 (3981, 3982)	otein sim. GBank gil4240287 dbj BAA74922.1 -	Contains protein domain (PF01602) - glycoprotein	glycoprotein	18108394, 56182575, 22278994, 35696286,
		(AB020706) KIAA0899 protein [Homo sapiens]	Adaptin N terminal region		56994075, 22278997, 22278999, 29331822,
_					29331824, 29331825, 60432289, 29331828,
					264508, 264906, 264907, 264908, 56182435,
					264510, 265007, 21906754, 33109954,
					87168474, 265017, 265018, 265019, 264762,
					18108351, 264763, 264683, 264369, 264288,
					264685, 264766, 264687, 264769, 21906765,
					21906768, 21906769, 55811957, 265020,
					60431528, 263974, 18108379, 35695855,
					264555, 264557, 264639, 83373044,
					18108384, 87158518, 60432113, 22279000,
100	_	Т	(200000) ciomak ciosaca aciosaca	, in the state of	222/9002, 254554, 254485 36234927, 264006, 264007, 264000, 265007
7661			Zing face Court August (Processy - Sunce	Since Since	25551021, 204500, 204507, 204509, 205007,
		lepeat on to [naposts sarconia-	Zilic iliiger, Oznz lype		204003, 204700, 204000, 204700, 21900700,
		associated retpesvirus]			204025, 204035, 204036, 10100363, 56526486, 264566, 264567
1993	$\overline{}$	80054763 (3985, 3986) Novel Protein sim. GBank gil2565091 (U80761) - CTG26		UNCLASSIFIED	264592, 35696423
	\neg	afternate open reading frame [Homo sapiens]			

6 0 15		. 25. 18.		¥.
18108394, 18108397, 35696286, 60424269, 29331827, 29331828, 35696052, 265006, 264512, 55811386, 265010, 265018, 265019, 25811150, 1810837, 264086, 56181562, 265020, 264081, 33657023, 264080, 5181562, 265020, 27486264, 18108370, 18108379, 35657109, 27486264, 18108370, 18108379, 3565855, 264534, 264634, 264635, 264555, 264555, 264551, 36182323, 18108382, 264563, 264564, 264565, 264555, 264566, 264565, 264565, 264566, 264566, 264567, 264566, 264567, 264566, 264564, 264566, 264565, 264564, 264566, 264564, 264566, 264565, 264564, 264566, 264564, 26464, 26464, 26464, 26464, 26464, 26464, 26464, 26464, 26464, 26464, 26464, 26464, 26464,	52646365, 52646842, 22278994, 22278995, 22278996, 35695286, 22278997, 22278998, 22278998, 262278997, 22278998, 22278998, 262278997, 22278998, 262278999, 262278999, 26227899, 26331827, 29331828, 3569652, 264106, 29331830, 52644045, 265007, 265017, 265019, 264761, 264369, 265017, 265019, 264761, 264369, 265017, 265019, 264761, 2644150, 265020, 265021, 60170615, 52644150, 2365726, 265021, 60170615, 52644150, 27486264, 3659563, 36596423, 3659565, 246631, 52644332, 56182323, 6017034, 83373044, 56526486, 22279002, 264566, 264656, 26466, 264657, 26466, 264667, 26464150, 27486264, 3659563, 36596423, 3669565, 264631, 52644332, 56182323, 264566, 264657, 26466, 26279002, 264566, 264657, 2652486, 22279002, 264566, 264657, 2652486, 22279002, 264566, 264657, 2652486, 22279002, 264566, 264657, 2652486, 22279002, 264566, 264657, 2652486, 22279002, 264566, 264657, 2652486, 22279002, 264566, 264657, 2652486, 22279002, 264566, 264657, 2652486, 22279002, 264567, 264667, 2652486, 22279002, 264567, 264667, 2652486, 22279002, 264567, 264667, 2652486, 22279002, 264567, 264667, 2652486, 22279002, 264665, 264657, 264667, 2652486, 22279002, 264666, 264657, 264667, 2652486, 22279002, 264665, 264657, 264667, 26562486, 22279002, 264665, 264657, 264667, 26567486, 22279002, 264667, 2656767, 264667, 2656764, 2656764, 2656764, 2656764, 2656764, 2656764, 266676, 264667, 266676, 264667, 266676, 264676, 264677, 26477, 26477, 26477, 264777, 264	264769, 52644229, 65274572, 21906768, 22278996, 35696286, 35695917, 255020, 22278999, 264534, 264490, 264259, 264692, 60432289, 33557109, 35696052, 264508, 264509, 18108374, 35696052, 264508, 264509, 264634, 264512, 265009, 264634, 264556, 264558, 264559, 264559, 264559, 264559, 264559, 264501, 264601, 264601, 264601, 264601, 264601, 264601, 264601, 264601, 264601, 264601, 264601, 264601, 264601, 264268, 264768	264907, 265006, 264910, 264603, 264692, 264629, 18108374, 264556, 264557	264905, 264906, 264908, 264910, 264596, 265017, 18108351, 264692, 264629, 264634, 264565
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	isomerase	UNCLASSIFIED
Contains protein domain (PF00515) - UNCLASSIFIED TPR Domain			Contains protein domain (PF00085) - isomerase Thioredoxin	
1999 94324903 (3997, 3998) Novel Protein sim. GBank gi[5225312 gb xAD40846.1 AF07244 - (AF072441) calcineurin binding protein cabin 1 [Homo sapiens]	95413705 (3999, 4000) Novel Protein sim. GBank gij1723232jspjQ10155jYATA_SCHPO - HYPOTHETICAL 90.6 KD PROTEIN C104.10 IN CHROMOSOME I	95072534 (4001, 4002) Novel Protein sim. GBank gi 107560 pir B38637 - Ras inhibitor (clone JC265) - human (fragment)	80236368 (4003, 4004) Novel Protein sim. GBank Gontains pre gij729433lsplP38657 ER60_BOVIN - PROBABLE PROTEIN Thioredoxin DISULFIDE ISOMERASE ER-60 PRECURSOR (ERP60) (58 KD MICROSOMAL PROTEIN) (PS9)	nover Protein sim. Gbank gilbb38tiprijlA27040 - neurofilament triplet M protein - chicken (fragment)
94324903 (3997, 3998)	95413705 (3999, 4000)	95072534 (4001, 4002)	80236368 (4003, 4004) 80074449 (4005, 4006)	00074449 (4003, 4000)
1999	2000	2001	2002	500

2004	95317318 (4007, 4008)	2004 95317318 (4007, 4008) Novel Protein sim. GBank gij4884249jembjCAB43230.1 - (/AL04998) hvpothetical protein Homo saniens	Contains protein domain (PF00076) - dna_ma_bind	dna_rna_bind	52645156, 52646842, 52646365, 56182575,
			RBD, or RNP domain)		35696286, 2227,0997, 2227,0999, 3227,0999, 60432049, 264259, 5245089, 2227,0999, 2227,0999, 2227,0999, 2227,0999, 264504, 29331824, 29331827, 35696052, 29331824, 29331826, 29331830, 264908, 264592, 6043336, 3365702, 264729, 264763, 26463, 265017, 265018, 265019, 264763, 264683, 264288, 2564429, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 265022, 25644150, 33657023, 52645129, 33657109, 33657182, 27486261, 27486262, 33657199, 27486265, 35696423, 25694332, 265621, 27486261, 27486262, 33657349, 27486265, 35696423, 35696423, 35696433, 35696433, 35696433, 35696433, 3654688, 87168518, 6143213, 264458, 87168518,
2002	87400864 (4009, 4010) Novel Protein sim. ((Z47812) similar to cDNA EST EMBL:C EST EMBL:D33965	Novel Protein sim. GBank gij3879501 emb CAA87795 - (Z47812) similar to ubiquitin carboxyl-terminal hydrolase; cDNA EST EMBL.D33366 comes from this gene; cDNA EST EMBL:D33965 comes from this gene; cDNA EST		ubiquitin	264488, 264906
0000		EMBL:D33822 comes from this gene; cDNA EST EMBL:D34547 comes from this ge			
2006		95351177 (40†1, 40†2) Novel Protein sim. GBank giļ4106673[emb]CAA22613] - (AL035064) queuine trna-ribosyltransferase [Schizosaccharomyces pombe]	Contains protein domain (PF01702) - UNCLASSIFIED Queuine tRNA-ribosyltransferase	NCLASSIFIED	56182575, 56181686, 22278995, 22278996, 22278997, 22278998, 22278999, 264259, 56182181, 60424269, 66714117, 35696052, 264906, 66712502, 264909, 264510, 6043336, 8565842, 256510, 255019, 265019, 26448, 264288, 264768, 29148627, 21906769, 2914874, 35695917, 60170615, 264691, 33657023, 65274620, 33657109, 55810764, 55811676, 35695855, 87168518, 6043713, 264663, 264482
2007	94325556 (4013, 4014)	94325566 (4013, 4014) Novel Protein sim. GBank gil26621611dbj BAA23712 - (AB007900) HH0452 cDNA clone for KIAA0440 has a 438-bp insertion at position 1711 of the sequence of KIAA0440. [Homo sapiens]	5	UNCLASSIFIED	264488, 263994, 35696052, 264508, 264905, 264906, 264907, 264909, 264909, 265007, 264901, 264909, 265007, 264910, 264592, 264758, 265011, 264760, 264762, 264766, 264766, 264769, 55811957, 3695917, 265020, 264691, 264693, 264628, 264632, 26433, 264634, 264637, 264637, 264638, 264637, 264638, 264637, 264638, 264637, 264638, 264637, 264638, 264648, 26
2008	85084428 (4015, 4016)	85084428 (4015, 4016) Novel Protein sim. GBank gij1550783 emb CAA69257 - (Y07960) homeodomain protein [Mus musculus]	Contains protein domain (PF00046) - homeobox Homeobox domain	omeobox	264909, 264768, 35695855

22278999, 264259, 264910, 264591, 265017, 264681, 264683, 21906768, 264691, 33657182, 3365748, 264631, 87168518, 264404, 22279002, 264563	52644507, 52845156, 52646365, 52646842, 18108397, 65274572, 22278994, 56994075, 35696286, 22278996, 22278994, 56994075, 35696286, 22278996, 22278999, 264259, 29331824, 29331824, 29331825, 29331826, 29331827, 29331824, 264311, 265007, 264512, 265009, 2650017, 265018, 265019, 264681, 264687, 265017, 265018, 265019, 264681, 264687, 265010, 264991, 264682, 265020, 5264429, 264691, 264682, 265020, 52644150, 264991, 264632, 264631, 264568, 264565, 264567, 264561, 264565, 264567, 264561,				ED 264488, 29331826, 264907, 264636, 264555, 264639, 264558
UNCLASSIFIED	ерћ	UNCLASSIFIED	UNCLASSIFI	dna_rna_bind	UNCLASSIFIED
	Contains protein domain (PF00057) - eph Low-density lipoprotein receptor domain class A		Contains protein domain (PF00409) - UNCLASSIFIED Kinesin light chain repeat	Contains protein domain (PF00096) - dna_rna_bind Zinc finger, C2H2 type	
2009 85749240 (4017, 4018) Novel Protein sim. GBank gij3882305 dbi BAA34512.1 - (AB018335) KIAA0792 protein [Homo sapiens]	95422458 (4019, 4020) Novel Protein sim. GBank gi 5262629 emb CAB45753.1 - (AL.080164) hypothetical protein [Homo sapiens]	9432B149 (4021, 4022) Novel Protein sim. GBank gil3347953 (AF076183) - cytosolic sorting protein PACS-1a [Rattus norvegicus]	87772137 (4023, 4024) Novel Protein sim. GBank gij1086678 (U41020) - coded for by C. elegans cDNA yk100g4.5; coded for by C. elegans cDNA yk100g4.3; weakly similar to human SREBP-2 basichelix-leucine zipper transcription factor [Caenorhabditis elegans]	94843842 (4025, 4026) Novel Protein sim. GBank gil4507985/refINP_003427.1 pZNF1 - zinc finger protein 135 zinc finger, C2H2 type (clone pHz-17)	87347940 (4027, 4028) Novel Protein sim. GBank gi[127720 sp P20938 MYP0_HETFR - MYELIN P0 PROTEIN PRECURSOR
85749240 (4017, 4018)		94328149 (4021, 4022) N	87772137 (4023, 4024) N	94843842 (4025, 4026) N	87347940 (4027, 4028) N
2009			2012	2013	2014

2015	88094922 (4029 4030)	2015 88094922 (4029 4030) Novel Protein sim CBank vilg 1286 hir 1822697 automin		COLUMN ACCION	E6107676 366060 364760 36606060
2	(200)	Volvox carteri (fragment)			20182373, 33030260, 204233, 330300032, 264508, 264906, 264907, 264510, 264512,
					87168474, 265010, 264681, 264288, 264689,
					264628, 35696423, 35695855, 264639,
					264563, 264564
2016	85298641 (4031, 4032) Novel Protein sim. (protein Tcp-10 - ma	Novel Protein sim. GBank gi 285046 pir S26413 - t-complex protein Tcp-10 - mouse		struct	264102, 264508, 264110, 265009, 33109954, 21906768, 265021, 33657109, 27486262,
	\rightarrow				263972, 18108374, 263976, 264555, 264564
2017	_			UNCLASSIFIED	264685, 264636
2018	79637067 (4035, 4036) Novel Protein sim. INVOLUCRIN	Novel Protein sim. GBank gi 124735 sp P18175 INVO_PIG-INVOLUCRIN			264693
2019	87787900 (4037, 4038) Novel Protein sim.	Novel Protein sim. GBank gi[2143910]pir S68216 -		phosphatase	264107, 264110, 264112, 265017, 263976
		phosphatase-1 glycogen-binding (GL)-chain - rat		•	
2020	94674476 (4039, 4040)	94674476 (4039, 4040) Novel Protein sim. GBank gil2078483 (U43200) - antifreeze		UNCLASSIFIED	264259, 264508, 264591, 265018, 264682,
		glycopeptide AFGP polyprotein precursor [Boreogadus saida]			264288, 264688, 22279002
2021	86718818 (4041 4042)	2021 86718818 (4041 4042) Naviel Protein eim CBack			20004075 204502 22400054 24000354
-	יייייייייייייייייייייייייייייייייייייי	AUTOCOOMIC SHIP COOLING			308940/3, 204393, 33109934, Z1900/34,
		gijosovajspjav odajer GM_ RAT - ELUNGATION FACTOR G, MITOCHONDRIAL PRECURSOR (MEF-G)			21906768, 33657023, 33657109, 27486261, 87168518
2022	95295665 (4043, 4044)	195295665 (4043, 4044) Novel Protein sim. GBank oil4218005 (AC006135) - putative			264757 264767 60170615 1R108385
		vicilin storage protein (globulin-like) [Arabidopsis thaliana]			
2023	2023 87722976 (4045, 4046) Novel Protein sim.	GBank	Contains protein domain (PF00442) - ubiquitin	ubiquitin	18108394, 22278999, 264259, 264905,
		gij5410230jgb AAD42992.1 AF07334 - (AF073344) ubiquitin Ubiquitin carboxyl-terminal	Ubiquitin carboxyl-terminal		264906, 264908, 264595, 264762, 264769.
		specific protease 3 [Homo sapiens]	hydrolases family 2		264634, 264636, 87168518, 60432113,
					22279000, 264482, 264565
2024	87896443 (4047, 4048)				60433438, 265017, 264686, 264692, 264693, 264636
2025	87858863 (4049, 4050) Novel Protein sim.	GBank	Contains protein domain (PF00637) - UNCLASSIFIED	UNCLASSIFIED	22278997, 264509, 264906, 264909.
		29055.1 AC00701 - (AC007018)	7-fold repeat in Clathrin and VPS		55812038, 265017, 265021, 265022,
		rabidopsis thaliana]			60170615, 264556
2026	94122114 (4051, 4052) Novel Protein sim. (Novel Protein sim. GBank gi 1655699 emb CAA69032 -		UNCLASSIFIED	56994075, 60432049, 264508, 66712502,
		(Y07752) pherophorin-S [Volvox carteri]			264112, 60170831, 87168559, 264288,
					264688, 264689, 21906766, 33657109,
					18108370, 264638, 18108385, 60432113,
					22279000, 22279002, 264564, 264566,
1000					264567
707	80249001 (4053, 4054)			UNCLASSIFIED	263978, 264634, 264486

2028	94699884 (4055 4056)	2028 94699884 (4055 4056) Novel Protein eim GBank gittgangat (1169766)			
		by C. elegans cDNA yk34b1.5; coded for by C. elegans	Contains protein domain (Process) - upiquitin IMECT-domain (ubiquitin-	uninbian	526445U/, 52645156, 52646842, 56182575, 56904075, 35696286, 22278997, 22278998
		cDNA yk13h10.5; coded for by C. elegans cDNA yk46e8.5:	_		2227275), 2233220), ELETOSSI, ELETOSSO, 22377, ELETOSSO, 22378000 60432040 264260 62646000
		coded for by C. elegans cDNA yk46d5.5; coded for by C.			29331822 29331824 66714117 29331825
		elegans cDNA yk43c2.5; coded for by C. elegans cDNA			29331826, 29331827, 29331828, 35696052
		yk46e8			264906, 264907, 264908, 29331830,
					52644045, 56182435, 265006, 265009,
					60432229, 33657402, 264595, 264757,
					55812038, 21906754, 52646317, 52644296,
					265010, 265011, 87168559, 265017, 265018,
					265019, 264448, 18108354, 264288, 264369,
					264766, 52644229, 21906765, 21906766,
					21906767, 21906768, 21906769, 55811957,
					35695917, 265020, 265021, 265022,
					52644150, 33657023, 65274620, 33657109.
					52645129, 18108368, 27486261, 27486262,
					27486264, 27486265, 35695763, 264629,
					55811576, 35696423, 35695855, 264635,
					264636, 52644332, 264558, 83373044,
0000	060000000000000000000000000000000000000				56526486, 22279000, 22279002, 264563
	32302032 (4027, 4028)	33302032 (4037, 4038) Novel Protein sim. GBank gij3599940 (AF017368) -	Contains protein domain (PF00621) - UNCLASSIFIED	UNCLASSIFIED	265009, 264595, 85658542, 264555, 264556,
$\overline{}$	2000 0000	faciogenital dysplasia protein 2 [Mus musculus]	RhoGEF domain		264557, 264558, 264559, 83373044
0502	91213/34 (4059, 4050)	91213/34 (4059, 4060) Novel Protein sim. GBank	Contains protein domain (PF00096) - transcriptfactor	transcriptfactor	18108394, 56994075, 22278997, 22278999,
		gil3530080jgbjAAD45825.1jAC00489 - (AC004890) similar	Zinc finger, C2H2 type		264259, 29331822, 29331824, 29331825,
		to HUB1; similar to BAA24380 (PID:g2789430) [Homo			66714117, 60432289, 29331826, 264108.
		[sapiens]			66712502, 264828, 265009, 265018, 265019,
					264681, 264682, 264684, 264685, 56181562,
					264689, 21906769, 265022, 264692, 264693,
					264628, 18108370, 18108374, 264634,
					264636, 264556, 18108385, 87168518,
Т					22279002, 264565
1502	60245281 (4051, 4052)				264591, 55811957, 18108365, 264557,
0000	10400 COOK TOOLOGO				264558, 18108382, 18108384
	91232607 (4063, 4064)	91232507 (4053, 4054) Novel Protein sim. GBank gi[5689491[db] BAA83029.1 -	Contains protein domain (PF00884) - hydrotase		65274572, 35696286, 29331824, 264908,
		(Abuzaudu) NiAA IU77 protein [Homo sapiens]	Sulfatase		265009, 264593, 265018, 264288, 264686,
					264769, 21906766, 21906767, 29148627,
					264628, 35696423, 264634, 264556,
					18108381, 60170394, 264559, 83373044,
2022	05000800 (4055 4055)	Control Of the Contro			18108385, 264482, 264484
	eannoons (4003, 4000)	appropriate (4005), 4000) Novel Protein Sim. Grank		synthase	56181562, 264628, 264632, 264555, 264556
		giz4840zolspl@640solcAG/_KAI - ALPHA-N-			
		SCHITCALACIOSAMINDE ALFRA-2,0- SIALYLTRANSFERASE (ST6GALNACIII) (STY)			
2034	91232529 (4067, 4068)	91232529 (4067, 4068) Novel Protein sim. GBank	Contains protein domain (PE00076) dos ros hind		22278006 22278000 254007 20224820
	•	gil4826984 ref NP_005147.1 pROD1 - UNKNOWN	RNA recognition motif. (a.k.a. RRM.		265008, 265018, 264681, 264682, 264684
			RBD, or RNP domain)		21906767, 21906768, 21906769, 33657109,
					83373044, 56526486

035	2035 83553451 (4069, 4070)	(264369, 264686, 265022, 56526486, 264567
<u>ş</u> [29331827, 29331828, 264682, 264369, 20148627, 60432113
2037		94324833 (4073, 4074) Novel Protein sim. GBank gil2734081 (AF000195) - similar		INCI ASSIETED	65274672 22278006 22278006 ECONOCIE
		to oxysterol-binding proteins (Caenorhahditis elegans)		dar issertation	052/45/2, 222/695, 222/8996, 569940/5,
		Supplied the supplied of the s			35696286, 22278997, 22278998, 22278999,
					264259, 29331824, 60432289, 29331826,
					29331828, 35696052, 264907, 29331830,
					66712502, 56182435, 265008, 265009,
					60170831, 264594, 55812038, 33109954,
					21906754, 87168559, 265017, 265018,
					265019, 264762, 264369, 264288, 21906765,
					21906767, 21906768, 21906769, 55811957,
					35695917, 265020, 265021, 265022,
					52644150, 33657023, 33657109, 33657182,
					35695763, 35695855, 264632, 264634,
					264636, 56182323, 83373044, 60432113,
2038	95422384 (4075 4078) Novel Bratis aim C	Maria Bratain aim OB I I O min aim aim			22279000, 22279002, 264563
3		//www.rividelisiii. dealik gijassubzajembjCABO/858j - //793785) predimed using Comfodor: similar to paka	Contains protein domain (PF01412) - UNCLASSIFIED	UNCLASSIFIED	22278995, 22278996, 56994075, 264259,
		(25020) predicted using Generalizer; Similar to KNA recognition most factor above a page of the page o	Putative GTP-ase activating protein		29331824, 35696052, 264905, 264906.
		FECT EMPLITATES SOME (SEE TRIM) TEST EMPLITATES	for Art		52644045, 265007, 265009, 87168559,
		ENDL. 10 1002 COMES IFOM THIS GENE; CUNA EST			265017, 18108351, 264448, 264369, 264766.
		EMBL: No.255 Curries from this gene; CDNA EST			264767, 264686, 18108358, 21906765,
		Liver. D27333 comes from this ge			21906769, 52644150, 33657023, 264692,
					18108362, 33657109, 27486262, 18108370,
					18108374, 18108379, 35696423, 65274791,
					264632, 264636, 18108383, 83373044,
					18108385, 87168518, 22279000, 22279002,
039	R5514628 (4077 4078)	2039 85514626 (4077 4078) Mariel Protein nim CBank All 2000 12 100 1 100 1			264563, 264564, 264566
3	לסיטר יייטר) סיטרייסטל	NOVEL FIGURE SIEE GEARK BILZZ4653[db] BAAZU813 -	Contains protein domain (PF00097) - UNCLASSIFIED	UNCLASSIFIED	22278997, 264259, 29331822, 264905,
			Zinc linger, C3HC4 type (RING		264906, 264907, 264908, 264909, 264510,
		•	linger)		265009, 264910, 264593, 264758, 265011,
					265018, 264762, 264288, 264766, 264768,
					264769, 21906766, 33657023, 264692,
					264693, 33657109, 35696423, 264631,
					264632, 264634, 264635, 264636, 264637,
	05308417 /A070 A0001				264639, 87168518, 264486
2043	05071736 (4079, 4000)	955004 I (4078, 4080)		UNCLASSIFIED	264592
	(7004 '1004) 05 (11006	-ingooggi - ingooggi -		rnapolymerase	264488, 22278998, 35696052, 264905,
		BILL SOUND STATE OF THE STATE O			264907, 264908, 264910, 265018, 264605,
		POLYMERASE 1135 KD POLYPEPTIDE (RNA			265019, 18108351, 264766, 264769,
		POLYMEKASE I SUBUNIT 2) (RPA135)			21906766, 265021, 265022, 264692,
					33657109, 264628, 264629, 35696423,
					35695855, 264637, 264638, 264563, 264564,
1					264565, 264567

2042	2042 95307447 (4083, 4084) Novel Protein sim. (AF131766) Similar sapiens]	Novel Protein sim. GBank gil4406590 gb AAD20040 - (AF131766) Similar to Ena-VASP like protein [Homo sapiens]	Contains protein domain (PF00568) - UNCLASSIFIED WH1 domain	UNCLASSIFIED	60424179, 35696286, 264259, 29331826, 35696052, 29331828, 264508, 264509, 264909, 264510, 264511, 265009, 264910, 264591, 60433356, 264591, 265019, 264681, 264764, 264369, 264765, 264684, 264288, 264769, 21900765, 35695917, 26453, 2544150, 264691, 254682, 18108356, 264558, 60170394, 18108365, 264558, 60170394, 18108365, 264568, 26279000, 22279000, 22279000, 22279000, 264482, 264563, 264564, 264566, 264664, 264664, 264664, 264664, 264664, 264664, 264664, 264664, 264667, 264462, 264566, 264664, 264664, 264667, 264462, 264669, 264666, 264667, 26467,
2043	94326076 (4085, 4086) Novel Protein sim. gil5052554lgblAAL BcDNA.GH06032 (Contains protein domain (PF00122) - transport	transport	264488, 52644507, 52646365, 56994075, 22278997, 22278997, 22278997, 22278997, 22281171, 264259, 29331824, 66714117, 29331826, 29331828, 33656970, 29146498, 264509, 264908, 5264045, 56182435, 265006, 33657402, 21906754, 52648296, 87168559, 265017, 265018, 256019, 264681, 264288, 265077, 265018, 256468, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 60170615, 264691, 33657023, 264683, 65274620, 33657109, 33657182, 27486261, 27486262, 33657499, 36695763, 18108374, 55811576, 35695855, 18108380, 18108374, 18108385, 56528486, 87168518, 5628486, 87168518, 5628486, 87168518, 5628486, 87168518, 5628486, 87168518, 5628486, 87168518, 5628486, 87168518, 5628486
	87106927 (4087, 4088) Novel Protein sim contains large com associated herpesv	Novel Protein sim. GBank gi 2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		struct	264093, 29331827, 264905, 66712502, 264592, 264689, 21906765, 21906769, 265020, 264692, 264482, 264566
2045					264692
2046		87320849 (4091, 4092) Novel Protein sim. GBank giļ4406698 gb AAD20062 - [(AF131852) Unknown [Homo sapiens]			264259, 264906, 264683, 22279002
2047		84578801 (4093, 4094) Novel Protein sim. GBank giļ4101720 (AF006466) - Iymphocyte specific formin related protein [Mus musculus]			22276999, 29147620, 29331824, 29146498, 264508, 265007, 265008, 265019, 264605, 265611, 29148627, 29148629, 265021, 33657023, 18108365, 33657109, 33657182, 18108377, 264556, 264638, 264559,
2048				UNCLASSIFIED	264909
	88094690 (4097, 4098) Novel Protein sim. (AB023223) KIAA1	Novel Protein sim. GBank gil4589656 dbj BAA76850.1 - (AB023223) KIAA1006 protein [Homo sapiens]		UNCLASSIFIED	264488, 264259, 29331824, 29331828, 35696052, 264906, 264907, 264908, 264909, 264910, 264603, 264763, 21906767, 21906768, 264629, 264634, 264637, 22279002, 264564, 264565, 264566, 264567

6	20077 20007				
	2030 / 9633635 (4099, 4100)			UNCLASSIFIED	264693
2051	2051 87780168 (4101, 4102)			UNCLASSIFIED	264488, 264259, 264509, 264906, 264907,
					264769, 18108374, 35696423, 264563,
					264566, 264486
202	88096393 (4103, 4104)	88096393 (4103, 4104) Novel Protein sim. GBank giļ4529889lgb AAD21812.1 -	Contains protein domain (PF00856) - kinase	kinase	264488, 263994, 35696052, 264508, 264905,
		(AF134726) G9A [Homo sapiens]	SET domain		264509, 264906, 264907, 264908, 264909,
					264113, 264511, 265009, 264910, 60170831,
					264592, 264758, 265010, 265011, 264605,
					264760, 264682, 264764, 264369, 264766,
					264686, 264768, 264769, 52644229, 264689,
					35695917, 33657023, 33657109, 264628,
					18108374, 35696423, 55811576, 35695855,
					264630, 264631, 264632, 264634, 264635,
					264636, 264556, 264638, 264639, 18108385,
					56526486, 60432113, 264563, 264564,
					264566, 264486, 264567, 264488, 263994,
					35696052, 264508, 264905, 264509, 264906,
					264907, 264908, 264909, 264113, 264511,
					265009, 264910, 60170831, 264592, 264758,
					265010, 265011, 264605, 264760, 264682.
					264764, 264369, 264766, 264686, 264768,
					264769, 52644229, 264689, 35695917,
					33657023, 33657109, 264628, 18108374,
					35696423, 55811576, 35695855, 264630,
					264631, 264632, 264634, 264635, 264636,
					264556, 264638, 264639, 18108385,
_					56526486, 60432113, 264563, 264564,
					264566, 264486, 264567
2023	2053 87763078 (4105, 4106) Novel Protein sim. G	Novel Protein sim. GBank gi[2995449]emb[CAA75113] -		UNCLASSIFIED	22278996, 22278997, 264259, 29331822,
		(114848) midline 1 protein (Mus musculus)			264102, 264508, 35695917, 263972, 264482
2054	95358937 (4107, 4108)	2054 95358937 (4107, 4108) Novel Protein sim. GBank gij3876326jembjCAB02090 -	Contains protein domain (PF00168) -		60424179, 264094, 264259, 29331825,
		(Z79754) similar to C2 domain [Caenorhabditis elegans]	C2 domain		60424269, 264906, 60432229, 60433356,
					87168559, 265019, 264760, 264288, 264686,
				•	21906769, 33657023, 264693, 55810764,
					55811576, 264635, 56182323, 60432113
2022	88259449 (4109, 4110)	88259449 (4109, 4110) Novel Protein sim. GBank		UNCLASSIFIED	264488, 29331826, 60432289, 29331828,
		gi 5353746 gb AAD42226.1 AF15913 - (AF159133) SIR2-			60433356, 265019, 264683, 264684, 265021,
		like protein [Oryza sativa subsp. indica]			33657109, 18108374, 264637, 18108385,
					87168518, 60432113, 22279000, 264564

<u>_</u>	88177396 (4111, 4112)	2056 88177396 (4111, 4112) Novel Protein sim. GBank	Contains protein domain (PF00749) - synthase	synthase	264488, 52645156, 56182575, 22278994,
		gij4826960jrefjNP_005042.1jpQARS - glutamine-tRNA	tRNA synthetases class I (E and Q)		35696286, 56994075, 22278996, 22278998,
_		synthetase			22278999, 60432049, 264259, 29331824,
					60432289, 29331827, 29331828, 33656970,
					264104, 264906, 264908, 265006, 265008,
					60170831, 264591, 60432229, 60433438,
_					18108348, 21906754, 33657084, 52644296,
					87168474, 265010, 87168559, 265017,
					265018, 264760, 18108351, 264681, 264682,
					264448, 264683, 264369, 264288, 264685,
					264687, 264688, 264689, 21906765,
					21906766, 21906767, 21906769, 55811957,
					35695917, 265022, 33657023, 18108362,
					33657109, 18108368, 33657182, 27486261,
					27486264, 27486265, 33657349, 264628,
					18108370, 264629, 18108374, 18108377.
					18108379, 35696423, 55811576, 20281152,
					264636, 264952, 18108385, 18108388,
					87168518, 264482, 264565, 264566, 264567
	87877905 (4113, 4114)	2057 87877905 (4113, 4114) Novel Protein sim. GBank		UNCLASSIFIED	52646842, 52646365, 56182575, 35696286,
		gij728850 sp P08640 AMYH_YEAST - GLUCOAMYLASE			22278996, 22278997, 22278998, 264093,
		S1/S2 PRECURSOR (GLUCAN 1,4-ALPHA-			52645080, 35696052, 29331828, 33656970,
		GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN			265009, 52646317, 55811386, 52644296,
		GLUCOHYDROLASE)			52644229, 21906769, 35695917, 265021,
					60170615, 52644150, 33657109, 33657182.
_					27486261, 27486262, 35695763, 35696423.
_	-				35695855, 52644332
	86276896 (4115, 4116)				265007, 265008, 264591
2059	79866684 (4117, 4118)	79866684 (4117, 4118) Novel Protein sim. GBank		UNCLASSIFIED	29331825, 264682, 264686, 264691, 264693,
		gi[119714 sp[P13983 EXTN_TOBAC - EXTENSIN Decembed Acet Matter Decembed the Dict			22279002
	- **-	GLYCOPROTEIN)			
2060	83050800 (4119, 4120)	83050800 (4119, 4120) Novel Protein sim. GBank gi 2811122 (U87318) - NaDC-2		UNCLASSIFIED	56182575, 29331824, 29331826, 264910, 55811957, 18108370, 55811576
1		Callopus rackis			3301.331, 10100310, 3301.313

2067	2067 95303892 (4133, 4134)				35696286, 22278997, 22278998, 60432049, 264259, 60432289, 60433438, 264682, 264448, 264369, 264288, 18108355, 21906765, 21906768, 265022, 33657109, 2656942, 3569428, 364588, 264404, 36486
2068	$\overline{}$			UNCLASSIFIED	264687
2069		94319177 (4137, 4136) Novel Protein sim. GBank gij3152662 (AF064604) - KE03	Contains protein domain (PF00023) - Iranscriptfactor Ank repeat	transcriptfactor	60424179, 56182575, 22278995, 22278996, 56994075, 264259, 29331822, 29331824, 29331825, 25331824, 29331825, 25331824, 28331825, 35696052, 29331828, 33558970, 284509, 264905, 56182435, 265009, 264504, 265019, 264448, 264764, 264766, 21906765, 21906765, 21906768, 21906769, 265021, 33657023, 33557109, 263976, 26455, 264557, 56182323, 33373044, 87168518, 60432113, 22279000, 22279002
2070	85791380 (4139, 4140)	85791380 (4139, 4140) Novel Protein sim. GBank gi]5712131 gb AAD47379.1 AF12049 - (AF120499) DEM1 protein [Homo sapiens]		UNCLASSIFIED	35695917, 264905, 264628, 264908, 264638
2071		86945116 (4141, 4142) Novel Protein sim. GBank gij3551531 dbj BAA33016 - (AB017437) avena [Gallus gallus]	Contains protein domain (PF00568) - WH1 domain		18108398, 265006, 265007, 265008, 265009, 264594, 265010, 265011, 18108351, 18108354, 18108368, 18108368, 18108388, 18108381, 18108385, 18108388, 18108391
2072	91718429 (4143, 4144) ,		Contains protein domain (PF00184) - Neurohypophysial hormones, C- terminal Domain		22278995, 35696286, 22278997, 22278998, 22278999, 264490, 60432049, 264259. 29331824, 29331824, 29331826, 3568052. 255008, 33657402, 21906754, 285011, 265019, 18108351, 254882, 264369, 21906766, 21906767, 21906768, 21906769, 35695417, 265020, 265021, 264659, 35696423, 2646555, 264556, 264558, 22279000
2073	27925664 (4145, 4146) Novel Protein sim. (D86976) similar to sapiens]	Novet Protein sim. GBank gil1504026 dbj BAA13212 - (D86976) similar to C.elegans protein (Z37093) [Homo sapiens]		UNCLASSIFIED	264556
2074	2074 94324767 (4147, 4148) Novel Protein sim. (AB020721) KIAAC	Novel Protein sim. GBank gi 4240317 db BAA74937.1 - (AB020721) KIAA0914 protein [Homo sapiens]			29331822, 264909, 264511, 265009, 264594, 264595, 264591, 265011, 265017, 265018, 265019, 264683, 265020, 265021, 18108370, 264632, 83373044, 264567

2085	94989476 (4169, 4170)	2085 94989476 (4169, 4170) Novel Protein sim GBank nit1655600 amblCAA60032	10:0:00	
		(Y07752) pherophorin-S [Volvox carteri]	UNCLASSIFIED	20182575, 60432289, 264908, 56182435, 87168474, 264763, 264369, 264686, 264693,
2086	91234404 (4171, 4172)	91234404 (4171, 4172) Novel Protein sim GRank ril/875032lembl/C4ABR0361	000000	18108370, 56182323
		(Z49125) similarity to Trichostronoulus paturations at the	UNCLASSIFIED	35696286, 264259, 35696052, 264906,
		Secretory brotein (Swiss Prot acression pumber D21037)		264907, 264908, 264909, 264910, 264759,
		CONA EST EMBL: 033349 comes from this neper: CONA		264604, 264762, 264768, 264769, 35695917,
_		EST EMBI 1037644 comes from this pens. CONA EST		263978, 35696423, 35695855, 264632,
		EMBL: D36149 come		264634, 264637, 264638, 264639, 56182323,
2087	21436337 (4173, 4174)			16108385, 264482, 264486
	94111527 (4175 4176)	94111527 (4175, 4176) Novel Protein eim CBank disapposolombic Adventage	UNCLASSIFIED	264489
		(AL021481) similar to Phosphoducomutase and		264488, 22278994, 35696286, 22278996,
		phosphomannomutase phosphoserine: cDNA EST		29331827, 3369052, 33657402, 21906754,
		EMBL: D36168 comes from this gene: cDNA FST		33109954, 87168474, 265017, 265018,
		EMBL:D70697 comes from this gene: cDNA FST vk373hg x		265019, 264448, 264683, 264369, 264685,
		comes from this gene, cDNA EST EMBL: T0080		2456272 24689, 21906765, 21906766,
				21906/04, 21906/06, 21906/69, 265020,
				205021, 205022, 204052, 3305/UZ3,
				3303/109, 3303/182, 2/486261, 2/486262,
				33657349, 27486265, 35696423, 35695855,
2089	95422801 (4177, 4178)	95422801 (4177, 4178) Novel Protein cim Chank		83373044, 87168518, 22279000, 264567
		GIA758118Irefine OGARS3 InDAP3 . Death accordade	cadherin	18108392, 264488, 52644507, 18108394,
		orizina a construction of the construction of		18108397, 52646842, 18108398, 56182575,
				22278994, 22278995, 35696286, 22278996,
				56994075, 22278997, 22278998, 22278999,
				264091, 264092, 264093, 264094, 60432049.
				264259, 29331822, 20281099, 29331824,
				29331825, 29331826, 29331827, 29331828,
				35696052, 33656970, 29146498, 29146499,
				264102, 264106, 264107, 264109, 264508,
_				264905, 264509, 264906, 264907, 264908,
				66712502, 264828, 52644045, 264909,
				56182435, 264110, 264112, 264510, 264511,
				265006, 264512, 265007, 265008, 264910,
				265009, 60170831, 264592, 264593,
				60433356, 33657402, 60433438, 264595,
				55812038, 264758, 21906754, 33657084,
				55811386, 52644296, 265010, 265011,
_				87168559, 265017, 265018, 265019, 264760,
				264761, 55811150, 264762, 18108351,
_				264682, 264448, 264763, 264764, 264683,
				264369, 18108354, 264288, 264685, 264766,
				264686, 264687, 264768, 52644229, 264688,
				18108358, 56181562, 264769, 18108359,
				264689, 21906765, 21906766, 21906767,
				21906768, 29148627, 21906769, 55811957,
				29148629, 29148784, 35695917, 265020,
				265021, 265022, 60170615, 264690,

6. 6. 17.	2. 768, 69.	59.	0, 6, 8, 528, 323,			77. 77. 7764. 2
22278995, 22278998, 22278999, 264259, 2931826, 35696052, 264910, 33657402, 60433438, 33109954, 8716844, 87168559, 265018, 265019, 264681, 264684, 264686, 264687, 264689, 21906765, 21906765, 21906766, 21906766, 21906765, 265022, 60170615, 33657023, 35696423, 35695855, 264952, 18108387, 22279000	263994, 264905, 264908, 264512, 265008, 264512, 265008, 264910, 55811386, 264288, 264768, 56181562, 21906765, 21906768, 21906769, 265022, 264628, 284563, 264567	22278997, 22278999, 66712502, 87168559, 264683, 265021, 264486	264905, 264906, 264907, 264908, 264510, 265006, 265007, 265009, 264910, 264596, 21906754, 87168474, 265011, 264603, 265018, 265019, 264760, 264766, 264768, 264769, 21906768, 21906768, 21906763, 265021, 264690, 33657023, 264693, 264623, 264634, 264634, 264637, 264637, 264557, 56182323, 2646564, 264636, 264637, 264557, 264557, 264557, 264557, 2645634, 264637, 264537, 264557, 264557, 264557, 2645634, 264564, 26466	65019,	35696286, 52844045, 265006, 265007, 265008, 87168559, 18108351, 21906769, 29148784, 265020, 33657023, 27486262, 18108374, 18108388	264488, 264489, 65274572, 56182575, 22278996, 22278997, 22278999, 264259, 60432289, 22278997, 22278999, 264259, 60432289, 264599, 264308, 264509, 264907, 264508, 2644045, 264909, 264500, 264907, 264512, 265009, 264909, 264510, 264512, 265009, 264500, 264512, 265009, 264500, 264512, 264008, 264764, 264607, 264608, 264764, 264608, 264764, 264689, 21906765, 21906766, 21906767, 264689, 21906765, 21906766, 21906767, 264699, 21906763, 264629, 264629, 264624, 36696423, 264639, 264596, 264566, 264489, 264565, 264566, 264566, 264566
3, 2227899 2, 264910, 1, 871684, 1681, 264, 1689, 219 1, 2190670, 2, 2190671, 1, 2190671, 1, 2190671,	1908, 264, 311386, 21 5, 2190670 1563, 264,	9, 667125(1486	1907, 264, 2609, 264, 265011, 265011, 1760, 264, 1906768, 257023, 264, 1637, 2647, 2	265018, 26	5, 265006, 18108351, 33657023,	774572, 56 7, 2227899 7, 3569604 1805, 264909, 26 1811, 265601 1806766, 1811, 1811, 1806766 1906766, 1811, 1811, 1806766 180672, 5, 1806766, 1806766 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 1806766, 180676,
22278998 35696052 33109954 5019, 264 4688, 264 21906767 170615, 3	4905, 264 4910, 558 21906765 4628, 264	22278999	4906, 264 5007, 268 87168474 5019, 264 906767, 2 4690, 336	265017.2	52644045 168559, 1 265020, 3 18108388	4489, 652 22278997 22331826 4509, 264 44045, 264 44063, 264 4762, 264 4288, 264 4288, 264 4289, 264 4489, 264 4486, 264 4486, 264
22278995, 22278998, 22278999, 264 29331826, 35696052, 264910, 3365; 60433438, 33109954, 87168474, 877 255018, 265019, 264681, 264684, 26 264687, 264689, 21906765, 21906766, 21906767, 21906769, 356 265022, 60170615, 33657023, 35698 35695855, 264952, 18108387, 22276	263994, 264905, 264908, 264511 265008, 264910, 55811386, 2642 56181562, 21906765, 21906768, 265022, 264628, 264563, 264567	22278997, 22278999, 667 264683, 265021, 264486	264905, 264906, 264907, 264908, 2647 265006, 265007, 265009, 264910, 2646 21906754, 87168474, 265011, 264603, 265018, 265019, 264760, 264766, 2647 264769, 21906767, 21906768, 219067 265021, 264690, 33657023, 264693, 26 264634, 264636, 264637, 264557, 5618	35696286, 265017, 265018, 265019, 18108388	35896286, 52844045, 265006, 265007 265008, 87168559, 18108351, 219067 29148784, 265020, 33657023, 274862 18108374, 18108388	264488, 264489, 65274572, 56182575, 2278996, 2278999, 26425 6093289, 2278999, 2278999, 2278999, 2278999, 2642803289, 2643899, 26278999, 264408, 264508, 264508, 264609, 264609, 264510, 264608, 264609, 264510, 264609, 264510, 264609, 264758, 264609, 264609, 264762, 264681, 18108351, 264689, 21906765, 21906766, 2190676, 264689, 21906765, 21906766, 2190676, 264689, 21906765, 21906766, 2190676, 264689, 21906765, 264629, 264629, 264629, 264629, 264639, 264565, 264483, 264556, 264567
		2 2	0000000			
	UNCLASSIFIED	нотеорох	rsport	ಲ	Contains protein domain (PF00035) - dna_ma_bind Double-stranded RNA binding motif	WD domain, G-beta repeat
	5	흔	3) - tran	MHC	5) - dne tif	O) - AT
			Contains protein domain (PF00153) - Irransport Mitochondrial carrier proteins		Contains protein domain (PF00035) Double-stranded RNA binding motif	eat (PF0040
			Contains protein domain (PFC Mitochondrial carrier proteins		n domain d RNA bi	beta repe
			ins protei		ns protei e-strande	ns protei
			Mitoch		Contai Double	Contai WD do
) camp		snb) CGI-40	mouse	. t:
	Novel Protein sim. GBank giJ4580997IgbJAAD24571.1JAF12108 - (AF121081) CAMP inducible 2 protein [Mus musculus]		Novel Protein sim. GBank gi[2352427 (AF004161) - peroxisomal Ca-dependent solute carrier [Oryctolagus cuniculus]	Novel Protein sim. GBank gil4929551[gb AAD34036.1 AF15179 - (AF151799) CGI-40 protein [Homo sapiens]	Bank gil1363238 pir A57284 - r RNA-binding protein Spnr - mouse	Novel Protein sim. GBank gi 3834423 (AF070689) - cytoplasmic dynein intermediate chain isoform DIC1a [Drosophila melanogaster]
	2108 - (A Is]		2427 (AF	5179 - (A	3238[pir] ng protei	chain isc
	Novel Protein sim. GBank gij4580997[gbjAAD24571.1]AF121 inducible 2 protein [Mus musculus]		nk gij235 ent solut	nk 36.1 AF1	nk gij136 NA-bindi	nk gij383
	im. GBa AAD2457 lein [Mus		im. GBa	im. GBal AAD3403 sapiens]	im. GBar ruclear R	im. GBai lanogast
	Protein s 0997 gb ble 2 prot		Protein s somal Ca ilus]	Novel Protein sim. GBa gi 4929551 gb AAD3403 protein [Homo sapiens]	Novel Protein sim. G spermatid perinuclea	Novel Protein sim. GBank cytoplasmic dynein interme [Drosophila melanogaster]
	Novel gij458 inducit		Novel	Novel gil492: proteir	sperm	Cytopi.
79, 4180	81, 4182	83, 4184,	85, 4186 8	37, 4188,	39, 4190,	91, 4192)
2090 88222470 (4179, 4180 <u>)</u>	95309161 (4181, 4182) Novel Protein sim. GBank gi 4580997 gb AAD24571. inducible 2 protein [Mus m	88223605 (4183, 4184)	87406073 (4185, 4186) Novel Protein sim. G peroxisomal Ca-depe cuniculus]	91230929 (4187, 4188) Novel Protein sim. GBank gil4929551gb AAD34036 protein [Homo sapiens]	95351526 (4189, 4190) Novel Protein sim. GBank gi 1363238 pir A57284 spermatid perinuclear RNA-binding protein Spnr	94119760 (4191, 4192) Novel Protein sim. GBank gil3834423 (AF070689) - Cytoplasmic dynein intermediate chain isoform DIC1 [Drosophila melanogaster]
2090	2091	2092	2093	2094	2095	2096

	264692,		264563			<u>. </u>	7,000	122	11828	64910	102,	174,	108351,	108355,	.68,		44	2006	:65010,			64908.	64760	64767.	74.	482,	0.5	100,	.020	17.		49	1576	2113,	
CENTACTO CLEASE CONTRACTOR	25274372, 204311, 203010, 204000, 205017, 264448, 264288, 265021, 60170615, 264692,	3365/109, 18108370, 264636, 264483	55994075, 264259, 264288, 265020, 264563		65274572 264905 65274444 264604	264636, 264555	56182575 22278995 22278996 22378997	22278998 60432049 264259 29331822	29331824, 29331825, 29331827, 29331828	29146498, 264909, 265008, 265009, 264910.	264591, 60432229, 60433356, 33657402,	264758, 21906754, 85658542, 87168474,	265017, 265018, 265019, 264681, 18108351,	264762, 264448, 264369, 264288, 18108355,	264686, 21906765, 21906767, 21906768,	21906769, 265020, 265021, 33657023,	18108374, 35696423, 264558, 83373044,	87168518, 60432113, 22279000, 2227	264091, 29331824, 264105, 265007, 265010,	25448B 264480 35505286 264250	204400, 204400, 330996200, 204239,	35596052,264508,264905,264907,264908, 264909 264511 264512 264501 264503	60433356, 264758, 264601, 264605, 264760	18108351, 264448, 264764, 264288, 264767,	264768, 21906769, 35695917, 18108374,	264634, 264555, 264559, 264563, 264482,	264486 65274572 56904075 22378990 264250	2011131E, 20231213, £££10333, £04233,	66712502, 265009, 60170831, 264595	33109954, 85658542, 87168559, 265017,	265019, 264448, 21906765, 21906768	265022, 33657023, 27486262, 33657349,	35695763, 60431528, 18108374, 55811576,	56182323, 18108387, 87168518, 60432113,	754654
		1 - 1 - 1 - 1	pnospnarase		phosphatase		nucl recot												UNCLASSIFIED	LINCI ASSIFIED										•					_
Contains protein domain (DE0006)	Zinc finger, C2H2 type	Contains assistant demand Appropriate	Dual specificity phosphatase	catalytic domain			Contains protein domain (PF00025) - nucl recot	ADP-ribosylation factor family																			Contains protein domain (PE01530) - Iranscriot/actor	Zinc finger, C2HC type						-	
2097 95322772 (4193, 4194) Novel Protein sim. GBank	.1pLYF1 - zinc finger protein,	87780340 (4195, 4196) Novel Protein sim. GBank	.1lpDUSP - dual specificity	ated)		byruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr IBos taurus	3881189 emb CAB16514 -			EMBL: C08337 comes from this gene; cONA EST	EMBL: CU9829 comes from this gene; cDNA EST yk291b4.5	Comes non mis gene; CDNA EST yk4						Novel Destroit aim OB and all 4500 400 at 100	(AB012808) mBOCT [Mus musculus]	87770461 (4203, 4204) Novel Protein sim. GBank gij3874149jembjCAA97423.1j -	(273103) predicted using Genefinder (Caenorhabditis	elegans]					74858.11 -								
95322772 (4193, 4194)		87780340 (4195, 4196)			95412927 (4197, 4198)		95332656 (4199, 4200)		_									87762604 (A201 A202)	(1502), 1502)	87770461 (4203, 4204)							95413576 (4205, 4206)								
2097		2098			2099		2100		_			_						2101		2102							2103								

					+	_		···	,	
264488, 56182575, 22278994, 56994075, 22278996, 22278996, 22278996, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22331824, 29331824, 29331826, 29331824, 29331826, 22331827, 29331828, 264508, 264905, 26369, 26407, 29331830, 265009, 60170831, 6043229, 33657402, 6043335, 264595, 6043249, 264768, 33557084, 81168474, 265010, 87168559, 266017, 265018, 265019, 264762, 18108351, 264684, 18108354, 264288, 264686, 21906766, 21906767, 21906767, 21906768, 21906769, 3565702, 264692, 33657023, 264693, 33657109, 264692, 33657023, 264639, 264558, 18108365, 264639, 264588, 18108385, 264639, 264588, 18108385, 264639, 264588, 18108385, 264639, 264588, 18108385, 264639, 264588, 18108385, 18108387, 565264886, 87168518, 60432113, 264564, 264568, 264567	265006, 265019	264906, 264639	264905, 264906, 264907, 264908, 264909, 264758, 265011, 264600, 264601, 264764, 264766, 264769, 264629, 3695855, 264632, 264634, 264635, 264638, 264638, 264639, 83373044, 264486	18108348, 264769, 18108370, 18108374, 264555, 264556, 264557, 264558	264564	264766, 35695917, 264630, 264567, 264486	264508, 264906, 264591, 264682, 22279002	265006	66714117, 29331826, 29331827, 60433438, 55812038, 265017, 265019, 284689, 21906769, 55811957, 265020, 265021, 33857109, 60170394, 284558	65274572, 264689, 264691, 264692, 60432113
UNCLASSIFIED			UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	struct	cadherin		UNCLASSIFIED
							Contains protein domain (PF00069) - struct Eukaryotic protein kinase domain	Contains protein domain (PF00028) - cadherin Cadherin		
2105 94848080 (4209, 4210) Novel Protein sim. GBank gi 1707032 (U80445) - coded for by C. elegans cDNA yk1395.3; coded for by C. elegans cDNA yk196.3; coded for by C. elegans cDNA CEMSE18F; coded for by C. elegans cDNA yk126b1.3; coded for by C. elegans cDNA yk65h8.3; coded for by C. elegans cDNA yk65h8.3; coded for by C. elegans cDNA yk65h8		Novel Protein sim. GBank gi 3881524 emb CAA93883 - (Z70038) ZK1067.4 [Caenorhabditis elegans]	94233976 (4215, 4216) Novel Protein sim. GBank gi 3176689 (AC003671) - Contains similarity to ubiquitin carboxyl-terminal hydrolase 14 gb Z35927 from S. cerevisiae. [Arabidopsis thaliana]		Novel Protein sim. GBank gil481043 pir S37671 - bat2 protein - human				gi[3327184 dbj BAA31660 - Nein [Homo saplens]	2115 90993785 (4229, 4230) Novel Protein sim. GBank gil4757890 ref nP_004328.1 pC8OR - chromosome 8 open reading frame 1
94848080 (4209, 4210)	2106 83365475 (4211, 4212)	79822662 (4213, 4214)	94233976 (4215, 4216)	80478719 (4217, 4218)		87818419 (4221, 4222)	87293783 (4223, 4224)	78941388 (4225, 4226)	87889342 (4227, 4228)	90993785 (4229, 4230)
2105	2106	-		2109	2110	2111	2112		2114	2115

2116 38226937 (4231, 4232) Novel Protein sin. GBank gil2246532 (J033672) - ORF 73.		25581624, 222.181939, 501821, 29.331824, 29.331824, 29.331827, 35696052, 264907, 56182435, 265008, 264591, 55812038, 55811386, 87168559, 264288, 264369, 29148629, 33657023, 35695763, 55811576, 35696423, 18103385	struct 29331824, 264511, 265009, 33109954, 265017, 265018, 264286, 264689, 265020, 264697, 56558486, 264487	glycoprotein 264259, 264905, 264907, 264908, 264510, 264511, 265009, 264910, 265010, 264602, 26428, 264768, 264693, 263967, 263972, 264568	UNCLASSIFIED 264693, 18108385	264091, 264259, 29331828, 29331828, 265020, 265017, 264601, 264288, 264685, 265020, 264691, 18108370, 55810764, 264555, 264636, 60432113	UNCLASSIFIED 264601, 264766, 263978	18108394, 56182575, 22278997, 29331822, 29331822, 29331828, 264907, 26182435, 265007, 264910, 265010, 265018, 264686, 265020, 55811576, 264555, 264937, 18108382, 83373044, 18108383, 18108384, 56526486, 264565, 264567			UNCLASSIFIED 264259, 264509, 264907, 264511, 85658542, 264763, 21906765, 35695917, 264636, 264486	UNCLASSIFIED 264489, 264489, 29331827, 35696052, 264906, 264908, 2649510, 265009, 264591, 264592, 264593, 33657402, 264594, 264595, 264596, 264769, 264601, 264603, 265018, 264604, 264604, 264603, 264761, 264603, 264761, 264603, 264764, 264604, 26	23657023, 33657109, 55810764, 264635,
86259387 (4231, 4232) Novel Protein sim. GBank gil2246532 (U93872) - ORF 7 Contains large complex repeat CR 73 (Kaposi's sarcoma associated herpesvirus)			sin-	sar	0	Contains protein domain (PF00017) - eph ing Src homology domain 2		40			ורא		
8659387 (4231, 4233) 1 1 1 1 1 1 1 1 1	Marie Bratain sim CBask ail2346632 // 102972 OBE 7	contains large complex repeat CR 73 (Kaposi's sarcoma associated herpesvirus)	Novel Protein sim. GBank gil2330021 (AF019250) - kine: elated protein; KRP; Costal2 [Drosophila melanogaster]	Novel Protein sim. GBank gi 1079307 pir B56573 - nuck oore complex glycoprotein p62 - African clawed frog	Novel Protein sim. GBank gi 4321407 gb AAD15748 - AF047690) ATP-binding cassette protein M-ABC1 [Homiapiens]	Novel Protein sim. GBank pil4885527 refINP_005480.1 pNSP3 - novel SH2-contain protein 3	Vovel Protein sim. GBank pil4757728IrefINP_004886.1 pAGTA - ingiotensin/vasopressin receptor AII/AVP-like	Novel Protein sim. GBank pil4929551Igb AAD34036.1 AF15179 - (AF151799) CGI- protein [Homo sapiens]	Novel Protein sim. GBank gi 2224551 db BAA20764 - AB002303) KIAA0305 [Homo sapiens]	Novel Protein sim. GBank gi 5689455 dbj BAA83011.1 - AB028982) KIAA1059 protein [Homo sapiens]	Novel Protein sim. GBank pil728831 sp P39188 ALU1_HUMAN - !!!! ALU SUBFAM! WARNING ENTRY !!!!	Novel Protein sim. GBank gil4539264 emb CAB39853.1 AL049495) conserved hypothetical protein Schizosaccharomyces pombe]	
المالية المالي	1 10FCA 14031 1A0301	- S - M	87788904 (4233, 4234)		86999317 (4237, 4238)	87789395 (4239, 4240)	80021375 (4241, 4242)	91230931 (4243, 4244)			95354041 (4249, 4250)		

35696286, 29331826, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264910, 265006, 264911, 265007, 265007, 265007, 265007, 264910, 264768, 265011, 264600, 264601, 264604, 264762, 264763, 264766, 264681, 264691, 264691, 264692, 264691, 264634, 264634, 264633, 264634, 264635, 264637, 264638, 264639, 18108385, 264563, 264564, 264563, 264563, 264563, 264563, 264564, 264563, 264563, 264563, 264563, 264563, 264563, 264563, 264564, 264563, 264563, 264563, 264564, 264563, 264563, 264564, 264564, 264563, 264563, 264563, 264564, 264563, 264564, 264564, 264564, 264563, 264564, 264564, 264564, 264564, 264567, 264563, 264564, 264567, 264563, 264564, 264564, 264567, 264563, 264564, 264564, 264567, 264564, 264567, 264563, 264564, 264567, 264564, 264567, 264564, 264567	56182575, 35696286, 56182181, 29331824, 60432289, 35696052, 264905, 264907, 66712502, 264909, 264909, 264907, 265009, 264910, 264591, 55812038, 265018, 264764, 264288, 264369, 264687, 264628, 264634, 264635, 264634, 264635, 264637, 56182323, 264639, 18108388, 264563, 264567	66714117, 264828, 264595, 55812038, 55811150, 55811957, 264693, 18108374, 263978, 65274791, 18108381, 83373044, 22279000	263981	35696052, 264909, 264768, 35695917	Contains protein domain (PF00122) - ATPase_associated 264488, 22278999, 264259, 29331827, 29331828, 35696052, 264509, 264905, 264905, 264906, 264907, 264906, 264907, 264906, 264907, 264906, 264910, 265007, 264910, 264510, 264506, 264517, 265007, 264910, 264510, 264506, 264595, 264760, 264591, 264591, 264691, 264591, 264691, 264591, 264691, 264691, 264691, 264691, 264691, 264691, 264760, 18108351, 264681, 264762, 264762, 264769, 264684, 264681, 264762, 264762, 264769, 264681, 264691, 264691, 264691, 264691, 264691, 264629, 264629, 264629, 264629, 264629, 264629, 264629, 264629, 264629, 264629, 264639, 264629, 264631, 264631, 264631, 264635, 264558, 264551, 264631, 264561, 2645
		UNCLASSIFIED	UNCLASSIFIED	potassium_channel	ATPase_associated
				Contains protein domain (PF00805) - Pentapeptide repeats (8 copies)	Contains protein domain (PF00122) - E1-E2 ATPase
2127 81118652 (4253, 4254) Novel Protein sim. GBank gil4868435[gb AAD31315.1 AF14323 - (AF143236) apoptosis related protein APR-2 [Homo sapiens]			95417144 (4259, 4260) Novel Protein sim. GBank gil2649255 (AE001012) - conserved hypothetical protein (Archaeoglobus fulgidus)	85723065 (4261, 4262) Novel Protein sim. GBank gi 1066886 (U41276) - Similar to Contains protein domain (PF00805) - potassium_channel 35696052, 264909, 264768, 35695917 potassium_channel protein. [Caenorhabditis elegans]	95361096 (4263, 4264) Novel Protein sim. GBank gil5689373JdbjgBAA82973.11 - (AB028944) KIAA1021 protein [Homo sapiens]
81118652 (4253, 4254)	87414262 (4255, 4256)		95417144 (4259, 4260)	85723065 (4261, 4262)	95361096 (4263, 4264)
2127	2128	2129	2130	2131	2132

FIED 60424179, 52646365, 52646842, 56894075, 35696286, 22278997, 22278998, 60432049, 56182181, 66714117, 60424269, 29331826, 29331828, 35696052, 264905, 264906, 264907, 66712502, 29331830, 56182435, 265006, 264512, 265008, 60431735, 6043335, 33657402, 55812038, 33109954, 21906754, 55811386, 265010, 264603, 265017, 265018, 265019, 55811150, 18108351, 264682, 24368, 2244229, 56181562, 21906766, 21906767, 21906768, 21906768, 21906776, 21906761, 256921, 60170615, 336590423, 65274791, 35696855, 264634, 60431850, 18108380, 56182323, 83373044, 18108385, 264566, 264566, 2249002, 264566, 264566, 264566, 264903, 22279002, 264565, 264566, 264903, 22279002, 264566, 264566, 264903, 22279002, 264566, 264566, 264903, 222779002, 264566, 264566, 264903, 222779002, 264566, 264566, 264903, 222779002, 264566, 264566, 264903, 222779002, 264566, 264566, 264903, 222779002, 264566, 264566, 264503, 264566, 264503, 264506, 264504, 264506, 26450		IFIED 22278999, 29331828, 35696052, 264906, 264908, 264910, 265009, 264591, 264458, 5264601, 87188559, 264601, 18108351, 264448, 264683, 264684, 264689, 3669763, 264684, 264682, 356957023, 264692, 356957023, 264631, 224635, 264636, 264637, 56182323, 264639, 22279002, 264654		IFIED 264569, 264909, 33109954, 264763, 21906768, 60170394, 18108385, 264563	ATPase_associated 264259, 29331828, 35696052, 264909, 265006, 265017, 265018, 18108351, 264288, 21906768, 33657023, 33657109, 264628, 18108374, 35695855, 264634, 264555, 264558, 264559, 264559	264905, 264910, 264591, 55812038, 55811386, 85658542, 264760, 18108351, 18108359, 55811957, 265020, 265021, 33657023, 18108364, 55811576, 83373044, 1810836, 65656486
UNCLASSIFIED		528) - UNCLASS	UNCLASSIFIED	UNCLASSIFIED	ATPase_a	
		Contains protein domain (PF00628) - UNCLASSIFIED PHD-finger				
2133 95351539 (4265, 4266) Novel Protein sim. GBank gil4220489 (AC006069) - hypothetical protein [Arabidopsis thaliana]	95412697 (4267, 4268) Novel Protein sim. GBank gij3675351 emb CAB09415 - (Z96047) DY3.6 [Caenorhabditis elegans]	88079813 (4269, 4270) Novel Protein sim. GBank gil5689559fdbjjBAA83063.1 - (AB029034) KIAA1111 protein [Homo sapiens]	84346479 (4271, 4272) Novel Protein sim. GBank gi[2662167 dbj BAA23715 - (AB007903) KIAA0443 [Homo sapiens]	87637716 (4273, 4274) Novel Protein sim. GBank gil4884110 emb CAB43262.1 - (AL050090) hypothetical protein [Homo sapiens]	Novel Protein sim. GBank gi[5174779]gb[AAD40696.1] - (U87804) 50 kDa protein [Caulobacter crescentus]	94843882 (4277, 4278) Novel Protein sim. GBank gi[3850821 emb CAA77135 - (Y18350) U2 snRNP auxiliary factor, large subunit [Nicotiana plumbaginifolia]
95351539 (4265, 4266)						
2133	2134	2135	2136	2137	2138	2139

2140	87645655 (4279, 4280)	2140 87645655 (4279, 4280) Novel Protein sim. GBank gil4417293lgblAAD204181 -		INCLASSIFIED	254488 254250 20221924 254404 254400
		(AC007019) unknown protein [Arabidopsis thaliana]			264509, 265006, 264759, 265018, 264448,
					264288, 21906768, 55811957, 265021,
					33657023, 27486265, 35696423, 264636,
3					264556, 264557, 264559, 264566
7 14 1				UNCLASSIFIED	265020, 264693
2142				UNCLASSIFIED	263978
2143	94140051 (4285, 4286) Novel Protein sim. G	Novel Protein sim. GBank gi 2135766 pir S53362 - mucin		UNCLASSIFIED	22278997, 29331827, 264907, 265020,
	2007	5AC (clone JER47) - human (fragment)			60432113
2144	94320114 (4287, 4288)	94320114 (4287, 4288) Novel Protein sim. GBank gi(2078483 (U43200) - antifreeze		UNCLASSIFIED	65274572, 264259, 29331824, 29331827,
		glycopeptide AFGP polyprotein precursor [Boreogadus			264906, 264908, 264591, 265011, 87168559,
		saida]			264600, 265019, 264288, 264768, 21906765,
					21906767, 55811576, 35696423, 65274791, 22279002
2145				UNCLASSIFIED	263978
2146		Novel Protein sim. GBank gi 1255871 (U53341) - short			264909 60433356 264686
		region of weak similarity to bovine membrane receptor p63			
24.47	C. COC. 10000	(Find Section I decide in a part of the section is a section in a sect			
214/	80432911 (4293, 4294)	80432911 (4293, 4294) Novel Protein sim. GBank gi 3080398 emb CAA18718.1 - {AL022603) putative protein [Arabidopsis thaliana]		UNCLASSIFIED	264907, 264768, 264769, 18108385
2148	80048811 (4295 4296)	80048811 (4295, 4296) Novel Profein sim CRank		0111004101411	00.11.00
		9i728837ispiP38194 ALU7_HUMAN - !!!! ALU SUBFAMILY SQ WARNING ENTRY !!!!		UNCLASSIFIED	264593
2149	87362022 (4297, 4298)	2149 [87362022 (4297, 4298) Novel Protein sim. GBank	Contains protein domain (PF00059) - alyconrotein		29331824 29331826 35696052 264758
		gij119863jsp P20693jFCE2_MOUSE - LOW AFFINITY IMMUNOGLOBULIN EPSILON FC RECEPTOR	Lectin C-type domain		87168474, 265018, 52644150, 33657109
		(LYMPHOCYTE IGE RECEPTOR) (FC-EPSILON-RII) (CD23)			
2150	94140059 (4299, 4300) Novel Protein sim. Gl	Novel Protein sim. GBank gi[5420387 emb[CAB46679.1 -		UNCLASSIFIED	22278998, 29331822, 29331824, 29331828,
		(AJ243459) proteophosphoglycan [Leishmania major]			264764, 264769, 21906766, 264486
1012	95353241 (4301, 4302)	95353241 (4301, 4302) Novel Protein sim. GBank gij5689407 dbj BAA82987.1 -			22278996, 56994075, 22278999, 60432049,
		(AB028958) KIAA1035 protein [Homo sapiens]			264259, 29331822, 29331824, 29331826,
					35696052, 29331828, 264508, 264511,
					60433356, 264758, 264596, 33109954,
				-	60174639, 265010, 265011, 87168559,
					265017, 265018, 265019, 264448, 264288.
					264689, 21906765, 21906766, 21906768,
					265020, 60170615, 33657109, 33657182,
					33657349, 18108370, 264635, 264557.
2162	70321640 (4303 4304)	Mount Destroy of the Charles and Charles and Charles			60170394, 18108385, 87168518, 22279000
3013	1905 1040 (4003, 4004)	rook lost (4303, 4304) Novel Froient Sim. Grank gij3452473 (AF084205) -		kinase	18108397, 18108398, 265007, 264591,
		serine/intenine protein kinase IAO1 [Kattus norvegicus]			265011, 18108351, 18108368, 18108374,
					18108388

2783	88313371 (4305, 4306)	2153 BB313371 (4305, 4306) Novel Protein sim. GBank gil4758704 ref NP_004216.1 pMASL - MFH-amplified sequences with leucine-rich tandem repeats 1	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat	glycoprotein	264488, 263994, 52646842, 22278996, 22278998, 22278999, 22278999, 264259, 29331822, 25696052, 264906, 264907, 264908, 264907, 264908, 264907, 264908, 264908, 56182435, 264510, 264511, 264512, 26458, 8716847, 264288, 264369, 264766, 264687, 264769, 52644229, 21906768, 21906768, 35695917, 33657023, 33657109, 35695855, 264631, 264632, 264642, 264642, 264642, 264642, 264642, 264642, 264642, 264642, 264642, 264642, 264642,
2154		Novel Protein sim. GBank gi[225150[prt] 1209265U - chorion protein 811 [Bombyx mori]		UNCLASSIFIED	56994075, 264094, 265009, 265019, 264288, 21906767, 35695917
2155	87424072 (4309, 4310)	·		UNCLASSIFIED	18108392, 18108398, 22278996, 264259, 29331824, 265008, 265010, 265011, 265017, 265019, 264288, 264686, 265020, 264693, 264528, 564823
2156	2156 64295205 (4311, 4312) Novel Protein sim. like; similar to AF0	Novel Protein sim. GBank gi[3970956 (AC004974) - spa-1- like; similar to AF026504 (PID:92555183) [Homo sapiens]			265007, 264684
2157	87316344 (4313, 4314)				264591
2158	86444218 (4315, 4316) Novel Protein sim. hypothetical protein	Novel Protein sim. GBank gij1076211 pirl SS0755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii		UNCLASSIFIED	264596
2159	80083729 (4317, 4318)		Contains protein domain (PF00651) - dna_rna_bind BTB/POZ domain	dna_rna_bind	29331822, 264112, 265009, 264691, 33657023, 264634
2160	16283674 (4319, 4320) Novel Protein sim. (AB007897) KIAAC				264634
2161	87739131 (4321, 4322)			UNCLASSIFIED	265008
2162	94319526 (4323, 4324) Novel Protein sim. (D86966) similarto			UNCLASSIFIED	65.74572, 264508, 264905, 264906, 264907, 264908, 5264049, 264909, 266907, 264907, 264909, 266507, 264907, 264591, 264591, 264593, 55812038, 264596, 264758, 265011, 264600, 264762, 264763, 264764, 264288, 264766, 264686, 264768, 264629, 263978, 264632, 264638, 264632, 264639, 18108385, 264563, 264565, 264567, 264567, 264636, 264563
2163	95417158 (4325, 4326) Novel Protein sim. (Z73974) cDNA ES CDNA EST yk291f elegans]	Novel Protein sim. GBank gij3876537]emb CAA98270 - (Z73974) cDNA EST yk291f5.3 comes from this gene; cDNA EST yk291f5.5 comes from this gene [Caenorhabdilis elegans]		UNCLASSIFIED	56182575, 22278996, 264093, 264683. 33657023, 65274620, 60432113
2164	80569456 (4327, 4328)		Contains protein domain (PF01006) - collagen Hepatitis C virus non-structural protein NS4a	collagen	264603, 264637, 264565

2165	94329169 (4329, 4330)	2165 94329169 (4329, 4330) Novel Protein sim GBank pil1086794 (1141107) - No		LINCLA COLETE	מממקבברת בממסלברבר שממסלברר שבמעמום
	•	definition line found [Caenorhabditis elegans]			30534015, £££16330, £££16331, £££16333, 26£259 29331822 29331824 20331825
					29331826, 29331827, 29331828, 264906
					29331830, 56182435, 265009, 21906754,
					33657084, 265011, 265019, 264448, 264288,
					264369, 21906765, 21906768, 21906769,
					265020, 265021, 264691, 264692, 33657023,
					65274620, 35695855, 264556, 60170394,
2166	87618034 (4334 4339)	87618034 (4334 4333) Novel Bestein Start Cont. 193055031			83373044, 60432113, 22279002, 264567
_	(2004, 1004) 4000	(VOSES) Ubiquiting attituting the control of the co		ubiquitin	52645156, 22278994, 22278998, 66714117,
		(1 15655) ubiquiun activating enzyme [Drosophila	_		29331828, 52644045, 265018, 265019,
		meianogasterj			264369, 21906765, 21906767, 21906768,
					21906769, 265021, 265022, 264693,
					27486262, 35695763, 18108376, 56526486,
9467	07746064 (4000 4004				87168518, 264567
	077 10004 (4555, 4554) NOVEL PROTEIN SIM. G	Novel Protein sim. GBank gij2224713[dbj BAA20840] -		UNCLASSIFIED	56182575, 35696286, 29331824, 29331826,
	_	לפוופולפים חווחים מספסים (בספססים)			29146498, 56182435, 265008, 265009,
					264592, 264593, 33657402, 33109954,
					265011, 265017, 265018, 18108351, 264369,
					21906764, 21906765, 21906768, 29148627,
					21906769, 52644150, 33657109, 35696423,
					18108381, 18108384, 18108385, 60432113,
	100000000000000000000000000000000000000				264567
0017	00333334 (4335, 4335)	oossesse (4555, 4555) Novel Protein sim. GBank gi 4321407 gb AAD15748 -	Contains protein domain (PF00664) - transport	transport	66714117, 29331827, 264907, 264511,
		(AF04/690) ATP-binding cassette protein M-ABC1 [Homo	ABC transporter transmembrane		264591, 265018, 264764, 264683, 264766,
\neg		sapiens	region.		264768, 264566
_	8/886937 (4337, 4338)				264629, 264555, 264559
2170	94141033 (4339, 4340) Novel Protein sim. G	Novel Protein sim. GBank		UNCLASSIFIED	65274572, 56182575, 22278997, 22278998,
					264259, 29331825, 264509, 264906,
		cotransporter KCC4 [Homo sapiens]			56182435, 60433438, 55812038, 264596,
					55811386, 265019, 264762, 264763, 264448,
					264764, 264684, 264288, 264766, 264685,
					56181562, 264689, 55811957, 265020,
					264535, 264691, 33657109, 60431528,
				-	18108374, 35696423, 55811576, 65274791,
					264634, 264639, 264558, 87168518,
т					60432113, 264564
7	60194050 (4341, 4342)			UNCLASSIFIED	264369, 265020, 264558
	85452460 (4343, 4344)				264259, 264558
2/2	8/035/40 (4345, 4346)	8/035/40 (4345, 4346) Novel Protein sim. GBank gi[4309681]gb[AAD15478] -		UNCLASSIFIED	264369
	20,00	(AC006930) R33423_1 [Homo sapiens]			
7	95003288 (4347, 4348)	93003288 (4347, 4348) Novel Protein sim. GBank aij2493778IsolO0456IYO35 CAFFI , PIITATIVE			264906, 35695855, 264555, 264557
	-	CUTICLE COLLAGEN C09G5.5			

IFIED	.05) - oncogene	kinase 18108392, 22278997, 22278999, 264093, 33657402, 265019, 264448, 264766, 264689, 21906768, 21906768, 21906769, 265021, 33657023, 18108370, 18108374, 60432113, 22279002)012) - leph 264488, 22278996, 22278999, 29331824,
	Contains protein domain (PF00805) - oncogene Pentapeptide repeats (8 copies)		Contains protein domain (PF000
	Novel Protein sim. GBank gij728837 spjp39194 ALU7_HUMAN - ii!! ALU SUBFAMILY Pentapeptide repeats (8 copies) SQ WARNING ENTRY i!!!	Novel Protein sim. GBank gi 5454072 ref NP_006416.1 pSLU7 - step II splicing factor SLU7	Novel Protein sim. GBank gil473407 (U08215) - NST-1 [Mus Contains protein domain (PF00012) - leph
2175 94325850 (4349, 4350) Novel Protein sim. [Araneus diadema	88223392 (4351, 4352) Novel Protein sim. -gij728837 sp P391 SQ WARNING EN	94128942 (4353, 4354) Novel Protein sim. 9i 5454072 ref NP. SLU7	87601557 (4355, 4356) Novel Protein sim
2175	2176	2177	2178

of (4359, 4360)	2180 95351397 (4359, 4360) Novel Protein sim. GBank gij31223171sp P90648 KMHB_DICDI - MYOSIN HEAVY CHAIN KINASE B (MHCK B)	Contains protein domain (PF00400) - kinase WD domain, G-beta repeat	kinas e	52644507, 22278994, 35696286, 22278997, 22778999, 264259, 52645080, 29331822, 29331824, 29331826, 29331827, 29331824, 29331826, 29331826, 29331826, 29331830, 264508, 264506, 265007, 26357402, 55812038, 21906754, 87168474, 87168559, 265017, 265018, 265019, 264763, 264682, 264683, 264684, 264288, 264686, 21906765, 21906769, 265020,
4362)	85764930 (4361, 4362) Novel Protein sim. GBank		kinase	265021, 265022, 52644150, 33657023, 33657109, 27486265, 33657349, 18108374, 35696423, 35695855, 263981, 60170394, 18108385, 56526486, 87168518, 60432113, 22279000, 264482, 264566, 264567, 264486, 29331827, 264369, 18108376, 264564
4364	gij3024689 sp Q15542 T2D4_HUMAN - TRANSCRIPTION INITIATION FACTOR TFIID 100 KD SUBUNIT (TAFII-100) (TAFII100)			
1, 4364)	87637731 (4363, 4364) Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	22278996, 22278997, 22278999, 264259, 29331822, 56182435, 264112, 264764, 264288, 21906767, 21906768, 21906769, 33657109, 18108376, 60170394, 22279000, 22279002
, 4366)	85460649 (4365, 4366) Novel Protein sim. GBank gij3873406[gb]AAC77482.1 - (U17129) unknown [Rhodococcus erythropolis]			264760
7, 4368)	87760690 (4367, 4368) Novel Protein sim. GBank gi 3114713 (AF061346) - Edp1 protein [Mus musculus]		tnf	29331822, 29331825, 29331826, 56182435, 265011, 264685, 264686, 21906768, 18108370, 264629, 264631, 264636, 264557
9, 4370)	87826453 (4369, 4370) Novel Protein sim. GBank gi 5106956 gb AAD39906.1 AF11361 - (AF113615) FH1/FH2 domain-containing protein FHOS (Homo sapiens)		-	29331824, 264907, 66712502, 264757, 265019, 264288, 264692, 56526486
1, 4372)		_	ATPase_associated	
87388173 (4373, 4374)			UNCLASSIFIED	35696052, 264905, 264907, 264907, 264908, 264511, 265006, 264910, 264758, 265019, 264762, 264681, 264766, 264769, 35695917, 264631, 18108388, 264566, 264486
5, 4376)	87771708 (4375, 4376) Novel Protein sim. GBank gil5107816 gb AAD40129.1 AF14941 - (AF149413) contains similarity to histone deacetylases; Pfam PF00850, Score=13.3, E=5e-10, N=1 [Arabidopsis thaliana]		histone	18108398, 56994075, 264259, 29331824, 29331825, 66714117, 29331827, 264906, 29331830, 265018, 265020, 265021, 56182323, 264559, 22279000, 22279002
7, 4378)	85693573 (4377, 4378) Novel Protein sim. GBank gi 3452357 (AF075724) - unknown [Legionella pneumophila]	Contains protein domain (PF01596) - O-methyltransferase		22278996, 264259, 29331826, 21906754, 264369, 264288, 263967

	8/639197 (4379, 4380)	2190 87639197 (4379, 4380) Novel Protein sim. GBank gil 132575[sp P29315 RINI_RAT - RIBONUCLEASE INHIBITOR		nucleaseinhib	22278996, 22278999, 29331822, 29331824, 29331826, 265008, 264910, 60170831, 55812038, 52644296, 265010, 265018, 264688, 56181562, 21906769,	
951	95198928 (4381, 4382)	Novel Protein sim. GBank gil5327002 emb CAB46272.1 - (Y18503) XAP-5-like protein [Homo sapiens]			35695917, 265022, 60170394, 22279000 29331825, 29331826, 29331830, 264510, 264511, 264910, 264593, 264594, 264556,	
=	126316 (4383, 4384)	1	Contains protein domain (PF00169) - PH domain		264559 264558	
<u>청</u>	(4385, 438 <u>6)</u>	94140073 (4385, 4386) Novel Protein sim. GBank gi[5420389Jemb CAB46680.1 - (AJ243460) proteophosphogiycan [Leishmania major]		UNCLASSIFIED	56181686, 29331825, 29331827, 264508, 264909, 265008, 264582, 60432229, 264288, 264684, 264786, 35695917, 33657023, 60431502, 60431528, 55810764, 55811576, 65274791, 35695855, 60431850, 56182323,	
21	418714 (4387, 4388)	21418714 (4387, 4388) Novel Protein sim. GBank gi[2773341 (AF040954) - putative protein phosphatase 1 nuclear targeting subunit [Rattus norvegicus]			60432113 264592	
88	083023 (4389, 4390)	Novel Protein sim. GBank gi 2832763 emb CAA15685.1 - (AL009191) /prediction=(method;; /prediction=(method;; /match=(desc;; /motif=(desc; [Drosophila melanogasier]		UNCLASSIFIED	22278996, 22278999, 35696052, 265006, 21906754, 265017, 35695917, 265021, 265022, 3569585	
,	031b31 (4391, 4392)	95091b51 (4391, 4392) Novel Protein sim. GBank gi 5262487 emb CAB45699.1 - (AL080076) hypothetical protein [Homo sapiens]		collagen	56182575, 35696286, 22278997, 22278999, 26429, 29331822, 66714117, 60432289, 29331827, 29331828, 264508, 265008, 265008, 265009, 60433438, 55812038, 265010, 265011, 264448, 264288, 264687, 5264687, 5264697, 265022, 264691, 33657023, 264691, 18108370, 18108376, 35696423, 55811576, 65274791, 35695855,	
950)73813 (4393, 4394)	95073813 (4393, 4394) Novel Protein sim. GBank gil4929567lgblAAD34044.1JAF15180 - (AF151807) CGL-49 protein [Homo sapiens]			264436, 56182323, 16103385 264768, 264769, 21906765, 21906766, 21906767, 29148627, 55811957, 35696286, 265020, 22278998, 255021, 264259, 33657023, 264693, 29331824, 35696052, 29331828, 18108370, 35695855, 264113, 265008, 264910, 60432229, 56182323, 33657402, 264768, 83373044, 21906754, 265018, 265019, 22279002, 264482, 264448,	
8	88060914 (4395, 4396)	Novel Protein sim. GBank gi[3548787 (AC005622) - R30953_1 [Homo sapiens]		UNCLASSIFIED	264565, 264288, 264369	

2199	88054355 (4397, 4398)	2199 88054355 (4397, 4398) Novel Protein sim. GBank gil2739372 (AC002505) -			264105, 264110, 264112, 264688, 55811957,
		hypothetical protein [Arabidopsis thaliana]			33657023, 264692, 263967, 20281071, 56526486
2200	87405385 (4399, 4400)	87405385 (4399, 4400) Novel Protein sim. GBank gij3043634 dbj BAA25481 - (AB011127) KIAA0555 protein [Homo sapiens]		struct	29331824, 264763, 264768
2201	94316872 (4401, 4402)	94316872 (4401, 4402) Novel Protein sim. GBank gi 3913470 sp 057314 DHBX_ANAPL - PUTATIVE STEROID DEHYDROGENASE SPM2	Contains protein domain (PF00106) - dehydrogenase short chain dehydrogenase	dehydrogenase	29331624, 35696052, 264905, 264907, 33657402, 55811386, 265017, 265018, 266019, 266019, 266020, 266020, 33657023, 3365709, 27486261, 18108370, 35696422, 3569585, 264556, 83373044, 87168518, 60432113
2202	91672385 (4403, 4404) Novel Protein sim. G (AL080186) hypothe	Novel Protein sim. GBank gi 5262665 emb CAB45767.1 - (AL080186) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264489, 264259, 29331824, 60432289, 35696052, 264905, 264909, 264592, 265017, 265018, 255019, 18188351, 264762, 264448, 264369, 264288, 264766, 21906765, 21906766, 264690, 264691, 264692, 33557109, 264534, 264636, 264558, 264639, 264558, 264559, 83373044, 18108385, 264404, 22279002, 264482
2203	87761832 (4405, 4406)	Novel Protein sim. GBank gi 1172845 sp P46629 RB25_RABIT - RAS-RELATED PROTEIN RAB-25	Contains protein domain (PF00071) - glycoprotein Ras family	glycoprotein	52646365, 56994075, 264259, 29331822, 29331826, 29331827, 29331828, 264910, 265010, 265011, 87168558, 265018, 265019, 264605, 264288, 21906769, 35695917, 33657023, 264692, 33657109, 3569563, 18108376, 264638, 22279000, 264566,
2204		88088671 (4407, 4408) Novel Protein sim. GBank gij121036]spjP29348]GBT3_RAT Contains protein domain (PF00503) - UNCLASSIFIED GUANINE NUCLEOTIDE-BINDING PROTEIN G(T), ALPHA G-protein alpha subunit (GUSTDUCIN ALPHA-3 CHAIN)	Contains protein domain (PF00503) -	UNCLASSIFIED	
2205	94147589 (4409, 4410)	94147589 (4409, 4410) Novel Protein sim. GBank gil4589480 dbj BAA76768.1}- (AB023141) KIAA0924 protein [Homo sapiens]	Contains protein domain (PF00096) - Idna_rna_bind Zinc finger, C2H2 type	dna_rna_bind	18108394, 18108397, 56182575, 60432049, 264259, 29331822, 29331826, 29331825, 29331826, 29331826, 29331826, 29331826, 29331827, 264906, 265007, 265008, 265009, 60432229, 265010, 265011, 265018, 24906768, 24906768, 21906769, 264690, 264691, 264693, 18108368, 55811876, 65274791, 294634, 18108381, 18108384, 60432113, 22279002, 264563, 264566
2206	20620008 (4411, 4412)	20620008 (4411, 4412)		UNCLASSIFIED	264591
2207	87787970 (4413, 4414)	Novel Protein sim. GBank gi 4557753 ref NP_000372.1 pMID1 - midline 1 protein	Contains protein domain (PF00622) - SPRY domain		29331822, 56182181, 29331827, 35696052, 52644045, 265006, 265019, 56181562, 55811957, 265021, 33657023, 35695763, 35695855, 60170394, 60432113, 264566
2208	86100830 (4415, 4416)				264906, 265019, 18108351, 21906769
5208	87800420 (4417, 4418) Novel Protein sim. ([Bos taurus]	Novel Protein sim. GBank gi 3986746 (AF105228) - tuftelin Bos taurus		struct	264112, 265009, 264691, 18108365, 18108374, 264634, 20281166

41 22 14	95361453 (4427, 4428)	95361453 (4427, 4428) Novel Protein sim. GBank gil4504325 reftNP_000173.1 pHADH - hydroxyacyl- Coenzyme A dehydrogenase/3-keloacyl-Coenzyme A thiolase/enoyl-Coenzyme A hydratase (trifunctional protein), alpha su	Ontains protein domain (PF00725) - dehydrogenase 3-hydroxyacyl-CoA dehydrogenase	dehydrogenase	264488, 52644507, 18108394, 56182575, 22278994, 22278994, 22278999, 264490, 66432649, 26278998, 22278999, 264490, 66432649, 2664286, 5654608, 29331822, 29147620, 29331824, 66714117, 29331825, 6613289, 29331827, 3569652, 29331828, 29331827, 3569652, 29331828, 20331827, 3569652, 29331828, 20331828, 20331827, 3569652, 264007, 265008, 26300, 60170831, 264591, 265007, 265008, 265009, 60170831, 265019, 265007, 265009, 264017, 265018, 265019, 18108351, 26448, 265010, 265011, 265011, 265011, 265012, 265019, 265012, 265013, 264687, 252644288, 264687, 21906768, 21906768, 21906768, 21906769, 33657023, 264690, 52644150, 265022, 264591, 33657023, 264690, 52644150, 265022, 264691, 33657023, 264692, 21408364, 33657023, 264639, 264639, 60431528, 33657623, 264638, 264639, 60431528, 3365562, 264638, 18108387, 18108388, 5652648, 87168381, 6043221, 33
2215		95419206 (4429, 4430) Novel Protein sim. GBank gil 1947160 (AF000298) - weak similarity to collagens; glycine- and proline-rich [Caenorhabditis elegans]			22279002, 264482, 264564, 264565, 264565, 264966, 264909, 265006, 264555, 264558, 87168518
2216		87614046 (4431, 4432) Novel Protein sim. GBank gil1572802 (U70854) - similar to Enterococcus faecalis TRAB (GI:388268) (Caenorhabditis elegans)		UNCLASSIFIED	264693
2217	80589404 (4433, 4434)	80589404 (4433, 4434) Novel Protein sim. GBank gi 5031707 ref NP_005503.1 pGARP - glycoprotein A repetitions predominant	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat	glycoprotein	264288, 33657109, 264556
2218		85518254 (4435, 4436) Novel Protein sim. GBank gij3878636jemb CA488953j - (Z49128) similar to cAMP-dependant protein kinase; cDNA EST EMBL:T00719 comes from this gene; cDNA EST yk46548.3 comes from this gene; cDNA EST yk46548.5 comes from this gene; cDNA EST yk492l4.3 comes from this gene; cDNA EST yw.	Contains protein domain (PF00069) - Eukaryotic protein kinase domain		35696423, 264563
2219		87614048 (4437, 4438) Novel Protein sim. GBank gil 1572802 (U70854) - similar to Enterococcus faecalis TRAB (GI:388268) [Caenorhabditis elegans]	Contains protein domain (PF01963) - TraB family		264682, 264683, 264688, 264689, 264693, 18108370, 18108376

		gil4507261/reflNP_003145.1pSTAT - statherin			264426, 18108394, 18108395, 35699286, 264426, 264097, 6043289, 264509, 264905, 264519, 264907, 265007, 264512, 264910, 2655009, 264591, 264510, 264511, 265007, 264512, 264910, 265009, 264594, 60433356, 264596, 265019, 264607, 264607, 264607, 264607, 264607, 264607, 264408, 264764, 264389, 264766, 264769, 264769, 264629, 264629, 18108362, 18108368, 264628, 264629, 18108362, 18108368, 264628, 264636, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 18108385, 264638, 264638, 18108388, 264638, 264638, 264638, 18108388, 264638, 264638, 264638, 18108388, 264638, 264638, 264638, 18108388, 264638, 264638, 264638, 18108388, 264638, 264638, 264638, 18108388, 264638, 264638, 264638, 18108388, 264638, 264638, 264638, 18108388, 264638, 264638, 264638, 18108388, 264638, 264638, 264638, 18108388, 264638, 264638, 264638, 18108388, 264638, 2646
2221	88060927 (4441, 4442)	88060927 (4441, 4442) Novel Protein sim. GBank gij3549154 (AC005625) - R27328 1 [Homo sapiens]			264483, 264566, 264486, 264567
2222				UNCLASSIFIED	264908, 265020, 35695855
2223				UNCLASSIFIED	265010, 264685, 264690, 264693, 264628, 263974, 263976, 55811576, 264555, 264638, 83373044, 264483
		87388515 (4447, 4448) Novel Protein sim. GBank gij3876005[emb CAA84799] - (Z35719) cDNA EST EMBL.D67419 comes from this gene; cDNA EST EMBL.C13853 comes from this gene; cDNA EST EMBL.C1378 comes from this gene; cDNA EST yk234a7.3 comes from this gene; cDNA EST yk234a7.5 comes from this gene; cDNA EST yk234a7.5	Contains protein domain (PF01958) - UNCLASSIFIED Domain of unknown function		264259, 264509, 56182435, 265006, 265008, 265008, 265009, 265009, 264693, 18108374, 18108385
2225	85749484 (4449, 4450)	85749484 (4449, 4450) Novel Protein sim. GBank gi 125847 (U53338) - C05E11.1 gene product [Caenorhabditis elegans]		transport	22278994, 22278995, 22278999, 52644045, 264600, 265019, 21906765, 21906769
	86978953 (4451, 4452)	Novel Protein sim. GBank gi[4826524 emb CAB42852.1 - (AL049848) hypothetical protein [Homo sapiens]			264259, 29331822, 29331824, 29331825, 29331827, 264508, 264906, 265007, 264691, 264634, 264486
2227	87721135 (4453, 4454)			UNCLASSIFIED	22278999, 265006, 265008, 18108354, 29148629, 29148784, 27486261, 18108374, 264637, 18108384
2228		91227337 (4455, 4456) Novel Protein sim. GBank gi 606976 (U16800) - ribonucleoprotein [Xenopus laevis]	Contains protein domain (PF00076) - dna_rna_bind RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	264091, 264092, 264094, 29331822, 29331825, 66714117, 264693, 263972, 264639, 83373044, 264563
2229		88060931 (4457, 4458) Novel Protein sim. GBank gij3549155 (AC005625) - R27328_2 [Homo sapiens]		UNCLASSIFIED	

264488, 264768, 52644507, 264769, 21906765, 21906766, 21906767, 21906769, 22278998, 3569591, 22278996, 22278999, 22278999, 25021, 264259, 5264129, 29331827, 264508, 264509, 264907, 18108370, 18108374, 35696423, 35695855, 265007, 264910, 224555, 33657402, 21906754, 18108387, 265010,	265018, 265019, 264760, 264288, 264567 264563	18108394, 56182575, 22278995, 35696286, 22278997, 22278999, 264259, 29331827, 35696052, 264907, 56182435, 265006, 265007, 265008, 264910, 264758, 55812038, 264603, 265018, 265019, 18108351, 264682, 264764, 264683, 264369, 29148629, 35695917, 264690, 5264150, 264691, 33657023, 264693, 18108374, 55811576, 266993, 18108374, 18108374, 55811576,	264906, 33657402, 265018, 264288, 264686, 2650000, 265000, 265000, 265000, 265000, 265000, 265000, 2650000, 2650000, 2650000, 2650000, 2650000, 2650000, 26500000, 26500000, 2650000, 2650000000000, 265000000000000000000000000000000000000	26020, 264693, 264688, 21906765, 265020, 264693, 264566, 65526486	22278997, 264563	264569, 264687, 264769, 265022, 264259, 60432049, 264691, 29331826, 60432289, 20281149, 264906, 264907, 264511, 265008, 265009, 264634, 284635, 264636, 264556, 264559, 264559, 60433438, 60432113, 264761,	264488, 264763, 294764 264488, 264769, 21906765, 21906766, 21906767, 21906768, 21906769, 22278995, 22278996, 22278997, 22278998, 265220, 265021, 22278999, 264259, 29331824, 29331826, 29331827, 35695763, 18108376, 35695855, 265007, 6043229, 33657402, 60433356, 60433438, 83373944, 18108385, 21906754, 18108387, 60432113, 22279000, 265019, 22279002, 264482, 18108351,
UNGLASSIFIED	UNCLASSIFIED	transcriptfactor		kinase	UNCLASSIFIED	de	kinase
					Contains protein domain (PF00400) - UNCLASSIFIED WD domain, G-beta repeat	Contains protein domain (PF00011) - eph Hsp20/alpha crystallin family	Contains protein domain (PF00400) - kinase WD domain, G-beta repeat
Novel Protein sim. GBank gij226154 pri 1412350A - DNA polymerase [Human adenovirus type 2]	88060937 (4461, 4462) Novel Protein sim. GBank gil3549154 (AC005625) - R27328_1 [Homo sapiens]	Novel Protein sim. GBank gi 5281316 gb AAD41476.1 AF13312 - (AF133124) transcription factor IIIC63 [Homo sapiens]	87755292 (4465, 4466) Novel Protein sim. GBank gij4249733jgb[AAD13780] - [AF109377] IdiBp [Mus musculus]	87771817 (4467, 4468) Novel Protein sim. GBank gil1706559jsplP54352[EAS_DROME - ETHANOLAMINE KINASE (EASILY SHOCKED PROTEIN)	91012316 (4469, 4470) Novel Protein sim. GBank gil4972734lgb AAD34762.1 - (AF132174) unknown [Drosophila melanogaster]	ocoos so 1 (4+7), 44 / 2) Novel Protein s.m. GBank gij 1082675 pir B53814 - p20	91012318 (4473, 4474) Novel Protein sim. GBank gil4972734 gb AAD34762.1 - (AF132174) unknown [Drosophila melanogaster]
2230 95342915 (4459, 4460) Novel Protein sim. G				87771817 (4467, 4468)	2235 91012316 (4469, 4470)		2237 91012318 (4473, 4474)

29331825, 265009, 264369, 33657109, 18108370, 18108374, 264557, 264559 264488, 65274572, 56182575, 35596286, 22276997, 22278999, 264259, 29331627, 35596052, 264508, 52644045, 56182435, 264511, 265007, 265008, 265009, 60433356, 60433438, 5581238, 21906754, 33657084, 55811386, 265018, 265019, 18108351, 264683, 264288, 264768, 264688,	4369, 33657109, 264557, 264559 264257, 23696286, 264259, 29331827, 644045, 55182435, 08, 265009, 6043335, 21906754, 33657084, 5019, 18108351, 68, 264687, 264688, 906768, 21906769, 5022, 60170615, 33657182, 33657349,	4369, 33657109, 264557, 2544559 182575, 35596286, 264259, 29331827, 644045, 56182435, 21906754, 33657084, 5019, 18108351, 68, 264687, 264688, 906768, 21906769, 5022, 60170615, 33657182, 33657349, 35696423, 35695855,	4369, 33657109, 264557, 264559 264259, 29331827, 264259, 29331827, 644045, 56182435, 08, 265009, 6043356, 21906754, 33657084, 5019, 18108351, 68, 264687, 264688, 906768, 21906769, 5022, 60170615, 33657182, 33657349, 33657182, 33657349, 33657182, 33657349, 33657182, 33657349, 33657623, 35695855,	4369, 33657109, 264557, 224559 264557, 224559 264259, 29331827, 264269, 29331827, 264045, 56182435, 21906754, 33657084, 5019, 18108351, 68, 264687, 264688, 906768, 21906769, 33657182, 33657349, 33657182, 33657349, 35696423, 35695855, 6754, 265010, 265011, 19, 264448, 264369, 265007, 21906769, 265007, 21906769, 265007, 265010, 265011, 264448, 264369, 3060766, 21906766, 21906769, 265020, 2486769, 2486760, 2486760, 2486760, 2486760, 2486760, 2486760, 2486760, 248676	4369, 33657109, 264559, 264557, 264559, 264559, 264559, 264559, 264559, 264559, 264559, 264559, 264569, 264045, 56182435, 08, 265009, 6043336, 21906754, 33657084, 5019, 18108351, 688, 26487, 264688, 9005768, 21906769, 33657182, 33657349, 33657182, 33657349, 33657182, 33657349, 33657182, 33696423, 35696423, 3569666, 265010, 265011, 265011, 265011, 265011, 265011, 265011, 26448, 264509, 265007, 264509, 265007, 264509, 265007, 264509, 265010, 265011, 26448, 264637, 264637, 264637, 264637, 264637, 264569, 26552489, 264637, 2645637, 2645637, 2645637, 2645637, 2645637, 264569, 26552489, 2645637, 2645637, 2645637, 2645637, 2645637, 264569, 26552489, 264637, 2645637, 2645637, 2645637, 2645637, 2645637, 2645637, 2645637, 2645637, 2645637, 2645637, 2645637, 26552489, 2645637, 264637, 2645648, 2645648, 2645648, 26457, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 2645648, 264	4369, 33657109, 264559, 2164559, 2564559, 264559, 264559, 264559, 264259, 29331627, 264045, 56182435, 08, 265009, 60433356, 21906754, 33657084, 5019, 18108351, 68, 264687, 264688, 2005768, 21906769, 250278997, 22278997, 22278997, 22278997, 22278997, 22278997, 22278997, 2505007, 265011, 265011, 265011, 265011, 265011, 19, 264448, 264369, 265020, 264637, 264	4369, 33657109, 264557, 264559 264257, 264559 264259, 29331827, 264269, 264259, 29331827, 264269, 26509, 60433356, 21906754, 33657084, 264687, 264688, 2006758, 21906769, 265278997, 22278997, 22278997, 22278997, 22278997, 22278997, 22278997, 256011, 19, 264481, 26439, 265011, 19, 264509, 265007, 265011, 265011, 265011, 264481, 264539, 265020, 265011, 264481, 264637, 26467, 26467, 2647,	4369, 33657109, 264557, 264559 264557, 264559 264259, 29331827, 664045, 56182435, 26182435, 21906754, 33657084, 5019, 18108351, 68, 264687, 264688, 9005768, 21906769, 252278997, 22278997, 22278997, 22278997, 22278997, 22278997, 22278997, 22278997, 22278997, 22278997, 264637, 264537, 2642896, 22278997, 26526480, 87168518, 26526480, 871680, 871680, 871680, 871880, 871880, 871880, 871880, 871880, 871880, 871880, 871880, 871880, 871880, 871880, 871880, 871880, 871880, 871880, 8	4369, 33657109, 264559, 22278996, 23696286, 246259, 29331827, 264425, 56182435, 246425, 56182435, 21906754, 33657084, 5019, 18108351, 68, 264687, 264688, 200768, 21906769, 21906769, 21906769, 21906769, 21906769, 21906769, 22278997, 22278997, 22278997, 22278997, 265010, 265011, 264011,	4369, 33657109, 264557, 224559, 226457, 224559, 264559, 264559, 264559, 264559, 264559, 264559, 264259, 29331827, 264688, 26906754, 33657084, 264688, 2009, 18108351, 2609, 18108351, 2609, 21906754, 33657349, 2005768, 21906769, 252278997, 22278997, 22278997, 22278996, 22278997, 264599, 264697, 264697, 264697, 264599, 264697, 264599, 2652488, 87168518, 26526488, 87168518, 871	4369, 33657109, 264557, 264559, 264557, 264559, 264559, 264559, 264559, 264559, 264559, 264559, 264559, 264559, 264559, 2646465, 56182435, 08, 265009, 60433366, 21906764, 33657469, 2006769, 21906769, 21906769, 21906769, 21906769, 21906769, 235696423, 35696855, 35696423, 35696423, 35696423, 35696423, 35696423, 3569643, 264637, 264637, 264637, 264637, 2655046, 87168518, 26550469, 2551196769, 2551196769, 2650463, 264637, 264637, 264637, 264637, 264637, 264638, 264768, 264638, 264768, 264638, 264638, 264638, 264638, 264638, 264638, 26504638, 264638, 264638, 264633, 26463463, 264634, 264634, 264634, 264644, 264644, 264644, 264644, 2646	22278996, 22278997, 22278996, 224687, 296459, 29331827, 264425, 56182435, 264425, 56182435, 264425, 56182435, 21906754, 33657084, 261906754, 33657084, 261906754, 33657084, 261906759, 18108351, 2622, 60170615, 23657182, 33657349, 235696423, 35695855, 22278997, 22278997, 22278996, 22278900, 222790000, 222790002, 2227900002, 222790000, 22279000, 22279000, 22279000, 22279000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 2227900000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222790000, 222700000, 222700000, 2227000000, 222700000000, 220700000, 22070000000, 2227000000, 2227000000, 222700000000
29331825, 265009, 264369, 33657109, 18108370, 18108374, 264557, 264559 264488, 65274572, 5618257, 35686286, 22278997, 22278999, 264259, 29331827, 3569625, 264508, 52644045, 56182435, 264511, 265007, 265008, 265009, 60433356, 60433438, 55812038, 21966754, 33657084, 55811386, 265018, 265019, 18108351, 264683, 264288, 264768, 2646887, 264688,	29331825, 265009, 264369, 33657109, 18108370, 18108374, 264557, 264559 264488, 55274879, 26182575, 35696286, 22278997, 22278999, 264258, 29331827, 3569629, 264258, 263336, 264511, 265007, 265008, 2564045, 56182435, 264511, 265007, 265008, 26905754, 33657084, 55811386, 265018, 265019, 18108351, 264683, 264769, 264768, 264687, 264688, 264769, 256027, 26502855, 26502855, 2650287, 265027, 265027, 265027, 265027, 26502855, 26502855, 265027, 265027, 265022, 26505855, 265027, 265027, 265022, 26505855, 265027, 265027, 265022, 26505855, 265027,	29331825, 265009, 264369 18108370, 18108374, 2245 2264488, 52274597, 261825 2278997, 22278999, 261825 22578997, 22578999, 2642 3569617, 265007, 265008, 2643343, 55812088, 21906 55811386, 265018, 265019 264683, 264288, 264768, 2 264769, 21906765, 2190677 35695917, 265021, 265022 52644160, 33657023, 3365 316955763, 18108370, 3569	29331825, 265009, 264369, 33657109, 18108370, 18108374, 284557, 264559 264488, 65274872, 56182575, 35696286, 22278997, 22278999, 264259, 29331826, 35596052, 264508, 52640465, 56182435, 264511, 265007, 265008, 264009, 264033, 264511, 265007, 265008, 21906764, 33657084, 55811386, 265018, 265019, 18108351, 264683, 264768, 21906768, 21906769, 33695917, 265027, 265022, 60170615, 35694160, 33657023, 33657182, 33657349, 35695781, 264693, 264699, 22278996, 22278996, 22278996, 22278996, 22278996, 22278996, 22278998, 264259, 29331827, 29331	29331825, 265009, 264369, 33657109, 18108370, 18108374, 264557, 264559, 264488, 65274572, 56182575, 2564559, 262488, 65274857, 2518995, 264259, 29331827, 3259999, 22278997, 22278995, 264259, 29331827, 3265908, 265009, 265009, 6043335, 6043348, 55812038, 21606754, 33657084, 55811386, 265019, 18108351, 264683, 265021, 265019, 18108351, 264683, 264288, 264768, 264688, 264769, 21906768, 21906769, 3569511, 265022, 60170615, 35695416, 33657182, 33657349, 35695418, 22278996, 22278996, 22278996, 22278996, 22278996, 265010, 264693, 265019, 264599, 265019, 265017, 265019, 264489, 265019, 264488, 265017, 265011, 265011, 265017, 265011, 265017, 265011, 265011, 265017, 265011, 265011, 265017, 265019, 2644289, 265010, 264488, 265010, 264488, 265010, 264488, 265010, 264488, 265010, 264488, 265010, 264488, 265010, 264488, 265010, 26488, 265010, 264488, 265010, 264488, 265010, 264488, 265010, 265011, 265010, 264488, 265020, 264289, 265009, 264500, 265010, 264017, 265011, 265010, 264488, 264509, 265000, 264501, 264488, 264509, 265000, 2645010, 264488, 264688, 264629, 265000, 264601, 2646	29331825, 265009, 264369, 33657109, 18108370, 18108374, 264557, 264488, 52278997, 22278999, 264259, 29331827, 25278997, 22278999, 264259, 29331827, 25278997, 22278999, 264259, 29331827, 264511, 265007, 265009, 264035, 264511, 265007, 265009, 264035, 264683, 264769, 18108351, 264683, 264769, 18108367, 264687, 264683, 264769, 1806769, 2650917, 265021, 265022, 60170615, 26504150, 33657023, 33657182, 33657399, 35695733, 18108370, 35696423, 35695855, 87168518, 22278996, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278995, 22278996, 22278997, 265009, 265007, 265010, 265010, 265010, 265010, 265010, 265010, 265010, 265010, 265010, 265010, 265013, 265010, 2650200, 2650200, 2650200, 2650200, 2650200, 2650200, 2650200, 2650200, 2650200, 2650200, 2650200,	29331825, 265009, 264369 18108370, 18108374, 2845 226488, 5227899, 261825 2267899, 2227899, 2642 35596052, 264508, 26440 264511, 265007, 265008, 2190 55811386, 265018, 265019 264683, 264768, 265019 264683, 264768, 265019 264683, 264769, 2190675, 219067 2564769, 21906765, 219067 2564593, 264629 22278996, 2227 264259, 29331822, 293318 22331827, 29331822, 293318 2265009, 264596, 21906754, 265019, 2	29331825, 265009, 264369 18108370, 18108374, 2445 264488, 65274572, 561825 22278997, 22278999, 2642 32596052, 264508, 26440 264511, 265007, 265008, 29044 264511, 265007, 265019, 26440 264769, 21906765, 219067 35695917, 265021, 26502 25644150, 33657023, 3365 35695917, 265021, 26502 256483, 264629 22278995, 22278900 26483, 264629 22278995, 222779000 264259, 29331822, 293318 22331827, 29331828, 2645 265009, 264966, 21906754, 265027 265009, 26496, 2190673 265007, 265018, 265019, 265027 265021, 33657109, 274862 265021, 33657109, 274862 265038	29331825, 265009, 264369, 33657109. 18108370, 18108374, 264557, 264559 264488, 65274672, 56182575, 35696286, 22278997, 22278999, 264259, 29331827, 3569625, 264508, 52644045, 56182435, 264511, 265007, 265008, 2564009, 604335 60433438, 55812038, 2190674, 33657084 55811386, 265018, 265019, 18108351, 264683, 264288, 264768, 264687, 264689, 35695917, 265021, 265022, 6170615, 35695917, 265021, 265022, 6170615, 35695617, 265021, 265022, 6170615, 35695618, 222799000 264693, 264696, 21906764, 29331826, 222789995, 22278996, 22278998, 265007, 264259, 29331822, 29331824, 29331826, 225009, 26496, 21906754, 265010, 26501 264269, 26496, 21906765, 21906766, 21906767, 21906768, 21906765, 21906766, 21906767, 21906768, 21906769, 264637, 265008 264504 18108374, 35695855, 284634, 264637, 265008 2645156, 22278995, 22278996, 22278997, 264504	29331825, 265009, 264369, 33657109, 18108370, 18108374, 264557, 264559 264488, 56274875, 56182575, 35698286, 22278997, 22278999, 264259, 29331827, 35596252, 264508, 52644045, 56182435, 264511, 265007, 265009, 264259, 29331827, 264511, 265007, 265009, 265009, 604335, 264769, 265009, 18106351, 264683, 264769, 21906769, 264769, 21906769, 26468, 264769, 21906769, 264683, 264769, 21906769, 264687, 265021, 265022, 60170615, 5264769, 21906762, 2031822, 26357899, 22278995, 22278996, 22278996, 22278996, 22278996, 22278996, 22278996, 22278996, 22278996, 22278996, 22278996, 22278997, 22278996, 22278996, 224448, 264369, 265007, 265019, 264448, 264369, 264488, 5264429, 21906764, 264637, 265021, 33657109, 27486262, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906766, 21906761, 36596855, 264634, 264637, 265021, 33657109, 27486262, 27486264, 18108374, 35695855, 284634, 264637, 264516, 22278995, 22278996, 22278996, 22278996, 22278999, 2331822, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 294491, 264512, 60433438, 264758, 264907, 264512, 60433438, 264758, 264604, 264512, 264512, 264604, 264612, 264604, 264612, 264612, 2646	29331825, 265009, 264369, 33657109. 18108370, 16108374, 264557, 264559 264488, 65274572, 5618257, 264559 26278997, 22278997, 2627699, 26429, 29331827, 3259697, 22278999, 264259, 29331827, 3269511, 265007, 265008, 265009, 6043336, 60433438, 56812038, 21906754, 33657084, 55811386, 265018, 265019, 18108351, 264683, 264788, 264768, 264688, 264683, 264789, 265020, 60170615, 35695717, 265021, 265022, 60170615, 35695763, 18108370, 35696423, 35695855, 87168518, 22279000 264593, 264629 22278995, 22278996, 265007, 264593, 29331828, 26509, 265017, 265009, 264596, 21906764, 265010, 265017, 265017, 265018, 265019, 26448, 264369, 264288, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265001, 36657109, 2748626, 27486264, 265006 264604 2626006 264604 22278999, 29331824, 29331824, 264504 262607, 264512, 6043348, 26478, 264604 21906776, 21906768, 29331824, 29331827, 264507, 264512, 6043348, 26478, 21906776, 21906769, 26448, 264604 262607, 264512, 6043348, 26478, 21906776, 21906769, 264764, 264687, 21906776, 21906778, 548478, 264604, 21906776, 21906778, 548478, 21906776, 21906778, 548478, 21906777, 264617, 264603, 264764, 21906776, 21906778, 549478, 21906777, 264617, 264603, 264764, 21906777, 264617, 264603, 264764, 21906777, 264617, 264603, 264764, 564687, 21906777, 21906778, 71906778, 21906777, 21906778, 71906778, 21906777, 264617, 264603, 264764, 264687, 21906777, 21906778, 719067789, 54841947	29331825, 265009, 264369, 33657109, 18108370, 18108374, 26457, 264559 264488, 65274572, 56182575, 35698286, 22278997, 22278999, 264299, 263295, 29331827, 35696286, 22644045, 55182435, 264511, 265007, 265009, 265009, 6043356, 6043348, 55812038, 21906754, 33657084, 55811386, 265018, 265019, 18108351, 264683, 264629, 33657182, 33657349, 35695418, 22278996, 22278996, 22278996, 22278996, 22278996, 265010, 265011, 265013, 264693, 264682, 21906765, 21906766, 21906765, 21906765, 21906766, 21906767, 21906768, 21906769, 264694, 56526486, 87168518, 264564, 264634, 264687, 264604, 264604, 264604, 264604, 264604, 264604, 264604, 264604, 264604, 264604, 264604, 264604, 264604, 264604, 264604, 264607, 21906768, 21906769, 264607, 21906768, 21906769, 264607, 264907, 2645126, 26278999, 29331822, 29331824, 29331827, 264907, 2645126, 22278996, 22278996, 22278999, 29331822, 29331824, 29331827, 264907, 2645156, 22278996, 25278996, 25278996, 25278996, 22278997, 264907, 2645126, 22278996, 2521906769, 264907, 2645126, 22278996, 2521906769, 264604, 264607, 21906768, 21906769, 264603, 264604, 264	29331825, 265009, 264369, 33657109, 18108370, 18108374, 264557, 264488, 22278997, 22278997, 22278997, 224508, 264259, 29331827, 25278997, 22278999, 264259, 29331827, 264511, 265007, 265009, 264045, 56182435, 264769, 18108351, 264769, 21906769, 18108351, 264769, 21906769, 264769, 21906769, 264769, 21906769, 264769, 21906769, 264769, 21906769, 264769, 21906769, 264769, 21906769, 264769, 21906769, 264687, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 265019, 264448, 264589, 265007, 265017, 265019, 264448, 264589, 265019, 264448, 264589, 265019, 265019, 265021, 265019, 26448, 264537, 265021, 33657109, 27486262, 21906765, 21906769, 265021, 265021, 265021, 265018, 264637, 265001, 264034, 264537, 265008, 264504, 264637, 265008, 264504, 264637, 265008, 264504, 264604, 264507, 265007, 264507, 265001, 264604, 264607, 265008, 21906769, 26448, 264681, 264504, 264507, 265008, 21906769, 29331824, 29331824, 29331827, 29331827, 29331824, 29331824, 264507, 265008, 21906769, 264604, 264507, 265001, 264604, 264507, 264607,
		SIFIED	SIFIED	SIFIED	SIFIED	SIFIED	SIFIED SIFIED					
Contains protein domain (PF00179) - ubiquitin Ubiquitin-conjugating enzyme	0179) - ubiquitin	0179) - ubiquitin UNCLASSIFIED	0179) -	0179) -	0179) -	0179) -	0179) -	0179) -	UNCLASSIFIEE Synthase UNCLASSIFIEE UNCLASSIFIEE UNCLASSIFIEE	UNCLASSIFIEE synthase UNCLASSIFIEE UNCLASSIFIEE UNCLASSIFIEE	UNCLASSIFIEE synthase UNCLASSIFIEE UNCLASSIFIEE UNCLASSIFIEE UNCLASSIFIEE	UNCLASSIFIEE synthase uNCLASSIFIEE UNCLASSIFIEE UNCLASSIFIEE
Ubiquitin-conjugating enzyme	njugating enzуme	njugating enzуте	njugating enzyme	njugating enzyme	njugating enzyme	njugating enzyme	njugating enzyme	njugating enzyme	Ubiquitin-conjugating enzyme Contains protein domain (PF0053	njugating enzyme	njugating enzyme	njugating enzyme
produkt unqukin-conjugating enzyme E2 [Picea markana]	Picea mariana)	Picea mariana)	Picea mariana) SLATION SUNIT (EIF-2B	SLATION SUNIT (EIF-28	SLATION 3UNIT (EIF-2B	SLATION 3UNIT (EIF-2B	SLATION SUNIT (EIF-2B 16417) - Similar elegans]					7
			S_RAT - TRANSL B GAMMA SUBUI	2. RAT - TRANSL B GAMMA SUBUI	S_RAT - TRANSL B GAMMA SUBUI	S_RAT - TRANSL B GAMMA SUBUI STOR)	Bank 41E2BG_RAT - TRANSLATION R EIF-2B GAMMA SUBUNIT (EIF-2B GE FACTOR)	2_RAT - TRANSL B GAMMA SUBUI TOR) [2291143 (AF016- Caenorhabditis ele [470340 (U00043)	E GAMMA SUBUI TOR) P291143 (AF016- Caenorhabditis et M70340 (U00043)	2_RAT - TRANSL B GAMMA SUBUI TOR) TOR) (2291143 (AF016- Caenorhabditis ele 1470340 (U00043) Seenorhabditis ele	2_EAT - TRANSL B GAMMA SUBUI :TOR) (2291143 (AF016 Caenomabditis ele 1470340 (U00043)	2_RAT - TRANSL B GAMMA SUBUI 3TOR) [2291143 (AF016- [470340 (U00043)] Aenorhabditis ele
			Novel Protein sim. GBank gi[2494312 sp P70541 E2BG_RAT - TRANSLATION INTIATION FACTOR EIF.2B GAMMA SUBUNIT (EIF-2B GDP-GTP EXCHANGE FACTOR)	Novel Protein sim. GBank gil2494312ispiP70541E2BG_RA- INITIATION FACTOR EIF-2B GAN GDP-GTP EXCHANGE FACTOR	ein sim. GBank (splp70541E2BG V FACTOR EIF-2E EXCHANGE FAC	in sim. GBank (splP70541E2BG V FACTOR EIF-2E EXCHANGE FAC	ein sim. GBank PACTOR EIF-26 EXCHANGE FAC:	Novel Protein sim. GBank gij2494312 sp 70541 E2BG_RAT - TRANSLATION gij2494312 sp 70541 E2BG_RAT - TRANSLATION INITATION FACTOR EIF-2B GAMMA SUBUNIT (EI GDP-GTP EXCHANGE FACTOR) Novel Protein sim. GBank gij2291143 (AF016417) - 3 to BZIP transcription factor [Caenorhabditis elegans] Novel Protein sim. GBank gij470340 (U00043) - similabeta-mannosyltransferase [Caenorhabditis elegans]	in sim. GBank IsplP70541E2BG V FACTOR EIF-2E EXCHANGE FAC in sim. GBank gil	ein sim. GBank IspIP70541[E2BG V FACTOR EIF-2E EXCHANGE FAC ein sim. GBank gil in sim. GBank gil in sim. GBank gil in sim. GBank gil	ein sim. GBank V FACTOR EIF-2B EXCHANGE FAC Exchange Fac in sim. GBank gil nscription factor [C ein sim. GBank gil sim. GBank gil sim. GBank gil	ein sim. GBank EXCHANGE FAC in sim. GBank gil
	6	[7]	2) Novel Protei (4) Novel Protei (9)(2494312); INITIATION GDP-GTP E	2) (A) Novel Protei (B) 12494312 k (B) 11471 ON (G) 1970 F	2) Novel Protei gil2494312 kint1ATION GDP-GTP E	1) Novel Protei gi[2494312]k initiation GDP-GTP E	1) Novel Protei gil2494312ki miTIATION GDP-GTP E	4) Novel Protei gil2494312 kim11ATION GDP-GTP E GDP-GTP E to BZIP tran Novel Protei to BZIP tran Novel Protei beta-manno:	1) Novel Protei gil2494312 gil2494312 iniTiATION GDP-GTP E GDP-GTP E to BZIP trani to BZIP trani to BZIP trani to BZIP trani beta-manno:	4) Novel Protei gil2494312k (NTTATION GDP-GTP E GDP-GTP E Deta-mannos	4) Novel Protei gil2494312k (Novel Protei to BZIP tran. (Deta-mannos) beta-mannos	4) Novel Protei gil2494312k (NTTATION GDP-GTP E GDP-GTP E Deta-mannos
	01051 14481 4480	80091951 (4481, 4482)	2241 80091951 (4481, 4482) 2242 91228075 (4483, 4484) Novel Protein sim. G 912443121splP705 INITIATION FACTO GDP-GTP EXCHAN	<u>91951 (4481, 4482)</u> 28075 (4483, 4484)	<u>91951 (4481, 4482)</u> 28075 (4483, 4484 <u>)</u>	91951 (4481, 4482) 28075 (4483, 4484) 02026 (4485, 4486)	2241 80091951 (4481, 4482) 2242 91228075 (4483, 4484) Novel Protein sim. G gi[2494312]sp P7056 iNITIATION FACTOI GDP-GTP EXCHANG 2243 78902026 (4485, 4486) 2244 85723527 (4487, 4488) Novel Protein sim. G to BZIP transcription	80091951 (4481, 4482) 91228075 (4483, 4484) Novel Protein sim. GBank 91228075 (4483, 4484) 91240312 8012016 8016 8016 8016 8016 8016 8016 8016 8	91951 (4481, 4482) 28075 (4483, 4484) 02026 (4485, 4486) 23527 (4487, 4488) 18545 (4489, 4490)	91951 (4481, 4482) 28075 (4483, 4484) 02026 (4485, 4486) 23527 (4487, 4488) 18545 (4489, 4490)	91951 (4481, 4482) 28075 (4483, 4484) 02026 (4485, 4486) 23527 (4487, 4488) 18545 (4489, 4490)	91951 (4481, 4482) 28075 (4483, 4484) 02026 (4485, 4486) 23527 (4487, 4486) 18545 (4489, 4490)
		2241 8009	2241 8005 2242 9122	2241 8006 2242 9123	2241 8009	2241 8009 2242 9122 2243 7890	2242 9122 2243 7890 2243 7890	2242 9122 2242 9122 2243 7890 2244 6572 2245 9531	2241 8009 2242 9122 2243 7890 2244 6572 2245 9531	2241 8009 2242 9122 2243 7890 2244 6572 2245 9531	2241 8009 2242 9122 2243 7890 2244 6572 2245 9931	2241 8009 2242 9122 2243 7890 2244 6572 2245 9531

2246	94848710 (4491, 4492)	2246 94848710 (4491, 4492) Novel Protein sim. GBank gild996096 db BAA78326.1 -	Contains protein domain (PF00153) - Iransport	transport	65274572, 22278995, 35696286, 22278996,
					22278397, 22278399, 204239, 33636032, 264106 264905 264907 265006 265007
					265008 60433438 33109954 87168559
					265018, 265019, 264288, 21906765,
					21906767, 21906768, 21906769, 55811957,
					35695917, 265020, 265022, 27486264,
					18108370, 18108374, 65274791, 35695855,
7567	7247 87862542 (4403 4404) Novel Destriction				60432113
j	JOSEPH (1180, 1181)	INDICATION TO SHIP OF SHIP SHIP SHIP STANDARD SHIP SHIP SHIP SHIP SHIP SHIP SHIP SHIP		UNCLASSIFIED	52645156, 52646365, 52645080, 35696052,
		(X83413) U88 [Human herpesvirus 6]			33656970, 52646317, 33657084, 265017,
					21906768, 21906769, 35695917, 33657109,
					52645129, 33657182, 27486261, 27486262,
	$\overline{}$				33657349, 27486265, 18108387
2248	95412996 (4495, 4496) Novel Protein sim.	GBank	Contains protein domain (PF00089) - cathepsin	cathepsin	264488, 264259, 264907, 29331830, 264909,
		02 ref NP_004123.1 pHABP - hyaluronan-binding	Trypsin		265007, 265009, 264595, 21906754,
		protein 2			65274444, 264603, 265019, 264762, 264448,
	-				264288, 264689, 21906766, 55811957,
					265021, 264691, 18108374, 264634, 264635,
					264636, 264555, 264638, 264557, 264558,
					264559, 18108383, 83373044, 18108385,
					264486
2249	94685662 (4497, 4498) Novel Protein sim.	GBank gij4038461 (AF107772) - TcST11		ebh	264766, 264628, 264636, 264637
		[Trypanosoma cruzi]	TPR Domain		
2250	79827508 (4499, 4500)	79827508 (4499, 4500) Novel Protein sim. GBank gij3738140 emb CAA21241 -		UNCLASSIFIED	264908, 18108374
		(AL031852) valyl-trna synthetase, mitochondrial precursor ISchizosaccharomyces nombel			
2251	87385863 (4501, 4502)	87385863 (4501, 4502) Novel Protein sim. GBank gij3218467lemb CAA07090.11-		UNCLASSIFIED	264259 35696052 264508 56182435
		(AJ006529) putative phosphatase [Gallus gailus]			265009, 264592, 264593, 264760, 264448.
					264684, 264288, 264690, 264628, 55811576,
					264555, 264556, 264557, 264558, 264559
					264566
2252	(87735867 (4503, 4504)	87735867 (4503, 4504) Novel Protein sim. GBank	Contains protein domain (PF01813) - synthase	synthase	264092, 264094, 264259, 29331822,
		gil4929325[gb AAD33953.1 AF14531 - (AF145316) vacuolar ATP synthase subunit D	ATP synthase subunit D		66714117, 29331828, 264102, 264103,
		proton pump delta polypeptide [Homo sapiens]			264104, 264105, 264109, 264112, 264511,
					265007, 60433356, 265010, 18108351,
					21906767, 21906768, 264691, 263974,
					263977, 264486, 264567
2253	91010703 (4505, 4506)		ו	UNCLASSIFIED	65274572, 265019

22564	95320031 (4507, 4508)	95320031 (4507, 4508) Novel Protein sim. GBank gil4502847 ref NP_001271.1 pCIRB - cold inducible RNA- binding protein	Contains protein domain (PF00076) - dna_ma_bind RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	264569, 18108394, 18108398, 56182575, 56994075, 36596286, 22278999, 264094, 60432049, 264259, 29331822, 29331824, 29331825, 29331824, 29331825, 29331827, 29331827, 29331826, 60432289, 289331827, 26331829, 264509, 264906, 264909, 264509, 265006, 264909, 264500, 265006, 264909, 264500, 265006, 26411, 265007, 265018, 26409, 264769, 264769, 265019, 265011, 87168559, 264288, 264289, 264766, 264686, 264768, 264691, 33657023, 264692, 18108370, 60431528, 264692, 18108372, 18108377, 18108379, 55811576, 35696423, 35695855, 264636, 264486, 264566, 264486, 264567, 18108385, 87168518, 60432113, 22279002, 264482, 264564, 264565, 264486,
2255		91010546 (4509, 4510) Novel Protein sim. GBank gil5541865[emb]CAB51072.1	Contains protein domain (PF00076) - struct RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	struct	65274572, 56182575, 22278997, 22278999, 264259, 29331822, 29331826, 29331826, 29331826, 29331826, 29331826, 29331826, 29331826, 29331826, 264508, 264905, 264907, 66712502, 264908, 264505, 2906754, 33109954, 87168474, 265011, 265017, 264682, 264764, 264768, 21906766, 21906766, 21906766, 21906766, 21906766, 21906769, 264690, 264696, 264908, 264636, 3657109, 264690, 264636, 3657109, 33657349, 264632, 264636, 52644332, 262279000, 22279002
2256		87020531 (4511, 4512) Novel Protein sim. GBank gij3327174 dbj BAA31655 - (AB014580) KIAA0680 protein [Homo sapiens]		UNCLASSIFIED	264768, 264689, 18108374
2257				UNCLASSIFIED	22278996, 22278999, 264681, 21906765, 21906766, 264567
2258		88090516 (4515, 4516) Novel Protein sim. GBank gij3025446 (AC004528) - R32184_2 [Homo sapiens]	Contains protein domain (PF00060) - misc_channel Ligand-gated ion channel	misc_channel	264908, 264592, 264764

18108396, 65274572, 56182575, 22278997, 22278998, 264259, 29331822, 29331827, 254905, 66712502, 264908, 264909, 56182435, 265007, 265008, 60432229, 33657084, 87168559, 18108351, 26448, 264683, 264389, 56181562, 265021, 60170615, 264690, 33657109, 60431528, 18108374, 52644332, 26182323, 18108385, 22279000, 22279000, 22278000, 2228800, 2228800, 222880000, 222880000, 222880000, 22288000, 222880000, 22280000, 22280000, 22280000, 22280000, 22280000, 22280000, 22280000, 22280000, 2228000000, 22280000, 222800000, 222800000, 222800000, 222800000, 222800000, 222800000, 22280000, 22280000, 22280000, 22280000000, 222800	56182575, 265020, 264905, 264906, 264908, 35696423, 264511, 264635, 55812038, 264758, 265018, 265019, 264605, 264760, 264563		56182575, 22278999, 29331822, 29331825, 60432289, 29331827, 35696052, 264508, 66712502, 5264045, 56182435, 265006, 265009, 265009, 6433336, 55812038, 265010, 265010, 265017, 265019, 264288, 264369, 721906765, 21906767, 55811957, 3565411576, 65274791, 56182323	22278994, 22278997, 264907, 264828, 52644150, 18108361, 264693, 18108374	264686, 264488, 264768, 264769, 264691, 264503, 264905, 264908, 264908, 264908, 264908, 264908, 264908, 265909, 264510, 264512, 265007, 265009, 264638, 264757, 264758, 18108395, 264566, 264486, 264766	264689, 264910, 264764	22278995, 22278997, 22278999, 264259, 265006, 265007, 265009, 60433438, 21906754, 265010, 265011, 265017, 26448, 26508, 264689, 21906768, 35695917, 265021, 18108374, 264638, 22279000, 22279002, 264566, 264487
UNCLASSIFIED	oxidase	UNCLASSIFIED			UNCLASSIFIED	UNCLASSIFIED	ribosomalprot
	Contains protein domain (PF00090) - oxidase Thrombospondin type 1 domain					Contains protein domain (PF00096) - UNCLASSIFIED Zinc finger, C2H2 type	Contains protein domain (PF01305) - ribosomalprot Ribosomal protein L15 amino terminal region
	Novel Protein sim. GBank gij3080663 (AC004614) - similar to f-spondin proteins AB006086 (PID:g2529225) [Homo sapiens]	Novel Protein sim. GBank gij3334526jembjCAA16138j - (AL021306) predicted using FGENEH [Homo sapiens]	Novel Protein sim. GBank gil4877759[gb AAD31421.1 AF12444 - (AF124440) MAGE tumor antigen D1 [Homo sapiens]			Novel Protein sim. GBank giļ477072 pir A48018 - mucin 7 precursor, galivary - human	Novel Protein sim. GBank gij3150479 (AF067212) - partial CDS [Caenorhabditis elegans]
		88074157 (4521, 4522)	91639292 (4523, 4524)	87602495 (4525, 4526)		86918663 (4529, 4530)	87773458 (4531, 4532)
2259	2260	2261	2262	2263	2264	2265	2266

2267	87395838 (4533, 4534)	2267 87395838 (4533, 4534) Novel Protein sim. GBank gij3560229jembjCAA20697.1j - (AL031530) hypothetical protein (Schizosaccharomyces pombe]		UNCLASSIFIED	35696286, 264259, 29331824, 29331825, 35696052, 29331828, 264905, 264509, 264512, 265009, 264907, 264907, 264909, 264512, 265010, 265018, 264762, 264482, 264588, 264669, 264768, 264429, 3656917, 264691, 33657422, 36564229, 35695417, 264691, 33657023, 34657310, 3466731, 3466731
2268	85693867 (4535, 4536)	85693867 (4535, 4536) Novel Protein sim. GBank		cadherin	18108381, 87168518, 264566 264488, 264259, 264509, 264588, 265010,
000		gil728832 sp P39189 ALU2_HUMAN - IIII ALU SUBFAMILY SB WARNING ENTRY IIII			265017, 264766, 18108385, 264486
5269	88177977 (4537, 4538)	88177977 (4537, 4538) Novel Protein sim. GBank gij103418 pirj S17885 - TcD37		UNCLASSIFIED	56182575, 60432049, 265007, 265009.
		protein - nut ny (crosoprina melahogaster)			264591, 87168559, 264605, 18108351, 21906764, 265020, 264629, 60431528,
2270	80410327 (4539, 4540)				264763
_	91010392 (4541, 4542)			cy10450	264909, 56182435, 265008, 55812038.
					55811957, 33657023, 264693, 33657109, 55811754, 55811576, 56182323
2272	84208220 (4543, 4544)			UNCLASSIFIED	264905, 264908
	95014271 (4545, 4546)	95014271 (4545, 4546) Novel Protein sim. GBank gil4176370 (AC005058) - similar	Contains protein domain (PF00462) -		52645156, 22278996, 22278999, 60432049
		to calcium-independent phospholipase A2; similar to			264259, 29331822, 29331824, 29331825,
		Sandra Sandra (Cir. gaada Cir.) Sabrens			29331826, 29331827, 35696052, 264909,
·					265006, 264593, 60433438, 21906754,
					200010, 204089, 21900705, 21900706,
					21906/6/, 21906/69, 265021, 265022, 60170616 264601 22657022 264602
					00170015, 204091, 3305/025, 204095, 33657109, 27486264, 18108376, 35606423
	~				35695855 264630 52644332 264558
					56182323, 22279002
5274	91640217 (4547, 4548)	91640217 (4547, 4548) Novel Protein sim. GBank gi 1480112 emb CAA67961 -	Contains protein domain (PF00538) - histone	histone	52645156, 22278997, 22278999, 52645080,
		(X99642) HP1-BP74 protein [Mus musculus]	linker histone H1 and H5 family		29331824, 29331825, 29331826, 29331827,
					29331828, 264905, 264908, 52644045,
					264511, 265008, 265009, 60170831, 264591,
					21906754, 33109954, 265011, 265018,
					18108351, 264448, 264288, 264684, 264766,
					21906765, 21906766, 21906767, 21906768,
					52644150, 264693, 18108364, 35695763,
					18108374, 35696423, 264634, 264557,
					264638, 52644332, 83373044, 18108385,
2275	88082501 (4549, 4550) Novel Protein sim. GBa	Novel Protein sim. GBank gi 3165406 (AC004755) -	Contains protein domain (PF00122) - transport	transport	56526466, 67166518, 22279002
2020	144007447 (4554	los3/502_2 [Homo sapiens]	E1-E2 ATPase		
_	1128/44/ (4551, 4552)			UNCLASSIFIED	264555, 264556

2277	88084123 (4553, 4554)	2277 88084123 (4553, 4554) Novel Protein sim Ghank rij2880070 (AC 004143), similar Journal Access (AC 004143)	Contains are left description		
	-	to murine leucine-rich repeat protein; possible role in neural	Leucine Rich Repeat	giycoprotein	ZZZ/8999, 35696052, 265008, 265019, 264369, 265020, 265022, 55810764, 264404
		development by protein-protein interactions; 93% similarity to D49802 (PID:q1369906) [Homo sapiens]			22279002
2278		94133079 (4555, 4556) Novel Protein sim. GBank gi 2618702 (AC002510) -		UNCLASSIFIED	18108394, 22278997, 29331826, 60433356,
		diviowi: protein (Arabidopsis mailana)			60433438, 21906754, 265018, 33657023, 264630, 83373044, 264666
2279		80419375 (4557, 4558) Novel Protein sim. GBank		UNCLASSIFIED	264766, 264565
_		gil119714 sp P13983 EXTN_TOBAC - EXTENSIN			
		FRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN)			
528					265008, 33109954, 265010, 265019, 265020
1977		74928.1 -	Contains protein domain (PF00400) - transport	transport	264092, 264259, 29331822, 29331824
		(AB020712) KIAA0905 protein [Homo sapiens]	WD domain, G-beta repeat	•	29331826, 35696052, 264107, 264906,
					264909, 52644045, 265006, 33657402,
					60433356, 264758, 265011, 265019, 264681,
					264683, 264684, 264686, 21906765,
					21906767, 21906768, 21906769, 60170615,
					264690, 52644150, 18108362, 264692,
					18108368, 18108374, 263978, 264631,
					18108381, 264559, 18108385, 56526486,
2282	87602829 (4562 4564)	87602820 (4563 4564) Noval Bratain sim OB1 114507070 (1900.00)			22279000, 264566, 264567
	מיייים (יייים איייי)	nover Forein sini. Obdin giji 357070 (003840) -		UNCLASSIFIED	264488, 264259, 29331822, 29331824,
		ייסים סיים ווחוש ווחוש פווחא ווחוש פווחא פ			29331827, 29331828, 29331830, 33657402,
					60433438, 87168474, 265019, 18108351,
					21906767, 21906769, 55811957, 33657023,
					52645129, 33657109, 33657182, 27486262,
2283	_	95362386 (4565 4566) Novel Protein sim CBank			263972, 55811576, 87168518, 20281169
		Ail3405720les[Oo2666]V304 HILLIAM VXXXXXIIIII		UNCLASSIFIED	60424179, 56182575, 22278994, 35696286,
		PROTEIN KIAAASA (HAAASA)			22278997, 22278999, 29331822, 29331824,
					56182181, 29331825, 29331827, 35696052,
-					29146499, 264905, 66712502, 264908,
					265007, 265009, 60432229, 264593,
					60431735, 60433356, 33109954, 33657084,
					55811386, 87168474, 265010, 265011,
					265018, 265019, 55811150, 264683, 264369,
					264288, 264688, 21906765, 21906767,
					21906768, 29148627, 21906769, 55811957,
					265020, 265022, 33657182, 27486261,
					18108370, 264628, 18108374, 55810764,
					18108379, 55811576, 35696423, 35695855,
					264630, 60431850, 263981, 18108382,
					83373044, 18108385, 18108387, 60432113,
					22279000, 264482, 264567

60424179, 52644507, 18108394, 52646842, 22278994, 5269286, 22278997, 22278994, 2669286, 22278996, 22278997, 22278999, 264259, 60432049, 29331822, 29331824, 29331824, 29331826, 29331827, 56182435, 33657402, 60433438, 33109954, 21906754, 8558542, 87168559, 265018, 265019, 55811150, 264682, 264369, 21906765, 21906769, 33657023, 33657023, 27486262, 27486264, 27486265, 18108376, 55810764, 35695456, 60431850, 87168518, 60432113, 264482, 264564	35696052, 264905, 264907, 264908, 264909, 264512, 265009, 264910, 264595, 264760, 18108351, 264682, 264763, 264689, 264666, 26468, 264693, 264629, 35695855, 264631, 264634	35696286, 56182435, 60170831, 264591, 60432229, 264592, 264593, 264594, 264595, 55812038, 264596, 87168474, 35695917, 264692, 55811576, 264555, 264557	264682	22278995, 56994075, 22278997, 22278999, 264259, 60432289, 264508, 264512, 265008, 33657402, 265017, 265018, 265019. 18108351, 264448, 21906765, 21906766. 21906767, 21906768, 33657023, 264557, 22279000, 22279002	264683	18108394, 264907, 265006, 265009, 33109954, 52646317, 265010, 18108351, 264681, 264682, 18108370, 18108374, 18108374, 18108365
	collagen	kinase	UNCLASSIFIED	tm7	kinase UNCLASSIFIED	struct
			Contains protein domain (PF01391) - UNCLASSIFIED Collagen triple helix repeat (20 copies)		Contains protein domain (PF00400) - kinase WD domain, G-beta repeat Contains protein domain (PF00400) - UNCLASSIFIED	WD domain, G-beta repeat
2284 95414955 (4567, 4568) Novel Protein sim. GBank gi[2498797]sp[G64311]PNAD_MOUSE - PROTEIN N- TERMINAL ASPARAGINE AMIDOHYDROLASE (PROTEIN NH2-TERMINAL ASPARAGINE DEAMIDOASE) (NTN- AMIDASE) (PNAD) (PROTEIN NH2-TERMINAL ASPARAGINE AMIDOHYDROLASE) (PNAA)	87781484 (4569, 4570) Novel Protein sim. GBank gij3342234 (U93909) - nuclear antigen EBNA-1 (Cercopithecine herpesvirus 15)	87737825 (4571, 4572) Novel Protein sim. GBank gij3873414 (U00043) - similar to D. melanogaster trithorax protein [Caenorhabditis elegans]	Novel Protein sim. GBank gil630905 pirl S42731 - collagen alpha 1 chain - sea urchin (Hemicentrotus pulcherrimus) (fragment)	ein sim. G	88084133 (4577, 4578) Novel Protein sim. GBank gil2887497 (AC004144) - R34001_1 [Homo sapiens] 88084137 (4579, 4580) Novel Protein sim. GBank gil2887497 (AC004144) -	R34001_1 [Homo sapiens] 84295281 (4581, 4582) Novel Protein sim. GBank gi 3253120 (AC005175) - R31449_3 [Homo sapiens]
95414955 (4567, 4568)		+	7 82986696 (4573, 4574) Novel Protein sim. Gi alpha 1 chain - sea u (fragment)			
	2285	2286	2287	2288	2289	2291

56182575, 35696286, 56994075, 29331824, 29331825, 35696052, 56182435, 60433438, 55812038, 33109954, 87168474, 87168559, 265018, 18108351, 264763, 26448, 264369, 21906766, 21906767, 21906769, 265021, 265022, 33657023, 264693, 65274620, 33657199, 2748624, 254629, 55811576, 35695555, 56182323, 56528486, 87168518, 22279000, 284567	264488, 18108397, 35696286, 264092, 264259, 29331822, 29331826, 264906, 264908, 264511, 264512, 265009, 264910, 18108351, 264764, 264389, 2644288, 264685, 264631, 264631, 264637, 18108381, 56182323, 264639, 18108385, 264404, 264563, 264565	22278997, 22278998, 22278999, 264259, 29331822, 60432289, 29331828, 35696052, 265018, 264684, 284288, 264686, 21906765, 21906766, 21906766, 21906768, 21906769, 265021, 264692, 33657109, 18108378, 35696423, 35695655, 264634, 22279000, 22279002, 264563, 264586	22278996, 60432289, 264682, 264683. 264689, 18108374 263974, 263978	264488, 65274572, 56182575, 22278997, 22278999, 264259, 29331822, 29331824, 29331826, 29331822, 29331824, 29331826, 29331828, 35696052, 264907, 264908, 52644045, 56182435, 264112, 265006, 265007, 264910, 265009, 60433356, 33657402, 264509, 2645019, 2645019, 2645019, 2645019, 2645019, 264691, 33657109, 55811576, 56182323, 60170394, 83373044, 18108385, 56526486, 264564, 264486
UNCLASSIFIED	struct	nuclease	UNCLASSIFIED	transcriptfactor
	Contains protein domain (PF00560) - struct Leucine Rich Repeat	Contains protein domain (PF00929) - nuclease Exonuclease		
2292 94328834 (4583, 4584) Novel Protein sim. GBank gil4803672 emb CAB42643.1 - (AJ133769) nuclear transport receptor [Homo sapiens]	87759213 (4585, 4586) Novel Protein sim. GBank gij3252981 (AF068921) - Ras- binding protein SUR-8 [Mus musculus]	86693580 (4587, 4588), Novel Protein sim. GBank gi 2062680 (U88964) - HEM45 [Homo sapiens]		94321251 (4593, 4594) Novel Protein sim. GBank gi 5689501 dbj BAA83034.1 - (AB029005) K AA1082 protein [Homo sapiens]
94328834 (4583, 4584)				
22292	2293	2294	2295 2296	2297

60424179, 56181686, 22278995, 35696286, 22278996, 22278999, 224999, 264490, 264259, 28331824, 66714117. 60424269, 28696052, 28331824, 66714117. 60424269, 38696052, 28331824, 66714117. 50424285, 38696052, 28331828, 66712602, 56182455, 284510, 265006, 60433438, 21906764, 33109954, 55811386, 265010, 264682, 264683, 264288, 264684, 264686, 21906767, 29148629, 55811957, 29148784, 35695917, 265020, 18108362, 38657023, 18108374, 56810764, 35696423, 35695855, 264630, 264634, 60431850, 18108386, 60431213, 222799000, 264482, 264567, 264488	264369	264488, 22278996, 22278999, 264259, 29331824, 66714117, 35696052, 264509.	264905, 264906, 264907, 264908, 264909, 264508, 264010, 265009, 264758, 265010	87168559, 264600, 265018, 264760, 264762,	18108351, 264764, 264766, 264768, 264769,	21906766, 21906767, 35695917, 265021,	264691, 33657023, 35695763, 18108370,	18108374, 35696423, 35695855, 264631,	264636, 264638, 18108385, 22279002, 264563	264908, 264758, 265017, 21906765,	83373044, 264563	52644045, 265019, 264288, 33657023, 18108370, 18108385	264259, 60432049, 264907, 264909, 264910,	60432229, 33657402, 265011, 265018,	264762, 264448, 264769, 264637, 264638, 83373044, 264486	264259, 29331824, 21906767, 33657182,	33657349	65274572, 22278996, 264908, 265006, 21906769, 264691, 264486
collagen	UNCLASSIFIED									struct	1	UNCLASSIFIED	struct			UNCLASSIFIED		
													Contains protein domain (PF00047) - struct	Immunoglobulin domain				
Novel Protein sim. GBank gij3875051 emb CAB02849 - (Z81050) predicted using Genefinder; similar to collagen; cDNA EST EMBL:D65564 comes from this gene; cDNA EST EMBL:D69046 comes from this gene; cDNA EST yk366b12.3 comes from this gene; cDNA EST yk366b12.5 comes from this gene		94124346 (4599, 4600) Novel Protein sim. GBank gi[2443886 (AC002294) - Unknown protein [Arabidopsis thaliana]								91235725 (4601, 4602) Novel Protein sim. GBank gil2143637 pir 184505 - calcium-	dependent actin-binding protein - rat	88084141 (4603, 4604) Novel Protein sim. GBank gi 2887497 (AC004144) · R34001 1 Homo sapiens	94141439 (4605, 4606) Novel Protein sim. GBank gil4884194 emblCAB43220.1 -	(AL049946) hypothetical protein [Homo sapiens]		94840434 (4607, 4608) Novel Protein sim. GBank	gij2494162lsp Q10005 YRY1_CAEEL - HYPOTHETICAL 39.9 KD PROTEIN T15H9.1 IN CHROMOSOME II PRECURSOR	90935911 (4609, 4610) Novel Protein sim. GBank gil4972686[gb]AAD34738.1 - [(AF132150) unknown [Drosophila melanogaster]
95312207 (4595, 4596)	80193720 (4597, 4598)	94124346 (4599, 4600) N								91235725 (4601, 4602) h			94141439 (4605, 4606) N			94840434 (4607, 4608) h	J (7)	_
2298	2299	2300								2301	3	2302	2303			2304		2305

Protein Planna Lapteral Planna Lapteral Lapteral Planna Lapteral Planna Lapteral Planna Lapteral Planna Lapteral La	9	95334940 (4611 4612)	Novel Protein sim GBank	Contains assisted demais (DE00400)	Lincoppositor	CONTRACTOR SOUTHERN SOUTH SOUT
79415283 (4613, 4614) 79415283 (4613, 4614) 67606408 (4615, 4614) 79415283 (4613, 4614)	_		014020565lphla4D34043 114E15180 (AE151908) CC1 49		viilasci cocpiui	204400, KKI10333, KKI10330, 33030200,
Protein Protein Protein Sapients Protein Sapients	_		gitazaponigujaranostos, ijar ibilou - (Ar ibilouo) CGI-48	IVVU domain, G-beta repeat		22278997, 22278998, 22278999, 264259,
9557218 (4613, 4614) 6760409 (4615, 4614) 6760409 (4615, 4619) 6760409 (4615, 4619) 6760409 (4615, 4619) 6760409 (4615, 4619) 6760409 (4615, 4619) 6760409 (4615, 4619) 6760409 (4615, 4619) 6760409 (4619, 4620) 6760409 (4619, 4620) 6760409 (4619, 4620) 6760409 (4619, 4620) 6760409 (4619, 4620) 6760409 (4619, 4620) 67721189 (4621, 4622) Novel Protein sim. GBank gil2137337 pit 46281 - gene Contains protein domain (PF00013) transcriptifactor mCBP protein - mouse	_		protein [Homo sapiens]			29331822, 29331824, 29331825, 29331826,
78415283 (4613, 4614) 8760409 (4615, 4619) Novel Protein sim GBank gild786732 refNPL_004522.1 pMOCS - mot/bdenum gild786732 refNPL_004522.1 pMOCS - mot/bdenum contact synthesis 2 95357218 (4617, 4618) Novel Protein sim GBank gild78059 genaplcA817070 - C269420, and EST EMBL-072444 comes from this gene; cDNA EST EMBL-0772046 comes from this gene; cDNA EST EMBL-072040 comes from this gene; cDNA EST EMBL-072044 comes from this gene; cDNA. EST EMBL-072044 comes from this gene; cDNA. EST EMBL-072044 comes from this gene; cDNA. EST EMBL-072040 comes from this gene; cDNA. EST EMBL-072040 comes from this gene; cDNA. ENGLY 46201 Novel Protein sim. GBank gi[2137337 pii 48281 - gene Contains protein domain (PF00013) - transcriptfactor KH domain CGP protein - mouse	_					29331827, 29331828, 35696052, 264508,
70415283 (4613, 4614) 9766409 (4616, 4616) Novel Protein sim. GBank 977218 (4617, 4618) Novel Protein sim. GBank git373737 pit 46281 - gene. 97837218 (4617, 4618) Novel Protein sim. GBank git2137337 pit 46281 - gene Charles ST EMBL.D72046 comes from this gene. CDNA EST EMBL.D72090 comes from this gene. CDNA EST EMBL.D72						264905, 264907, 29331830, 264908, 264909,
78415283 (4613, 4614) 67606409 (4615, 4614) 67606409 (4615, 4615) Novel Protein sim. GBank gig137305 comes from this gene. cDNA EST EMBL. D7284 comes from this gene. cDNA EST EMBL. D7280 comes from this gene. cDNA EST EMBL. D72804 comes from	_					264510, 265006, 264511, 265007, 264512,
79415283 (4613, 4614) 87606409 (4615, 4614) 87606409 (4615, 4616) Novel Protein sim. GBank 914758732[reftNP_004522.1]pMOCS - motybdenum 95357218 (4617, 4618) Novel Protein sim. GBank gij378059[emb]CAB17070] - (299942), cDNA EST EMBL-D73444 comes from this gene; cDNA EST EMBL-D73030 comes from this gene; cDNA EST EMBL-D7304	_					265008, 265009, 33657402, 21906754,
78415283 (4613, 4614) 87606409 (4615, 4619) 87706409 (4615, 4619) 87706409 (4615, 4619) 877218 (4617, 4619) 877218 (4617, 4619) 877218 (4617, 4619) 877218 (4617, 4619) 87721189 (4621, 4620) 87721189 (4621, 4620) 87721189 (4621, 4622)	_					85658542, 265010, 265011, 264601, 265017,
78415283 (4613, 4614) 87606409 (4615, 4616) Novel Protein sim. GBank 97606409 (4615, 4616) Novel Protein sim. GBank 977873218 (4617, 4619) Novel Protein sim. GBank g103739399imblCAB17070] - 95357218 (4617, 4619) Novel Protein sim. GBank g10373937ipit 46281 - gene: DNA EST EMBL.D72306 comes from this gene: CDNA EST EMBL.D72306 comes from this gene: CDNA EST EMBL.D72300 comes from this gene: CDNA EST EMBL.D7244 comes from this gene: CDNA. 72601668 (4619, 4620) 87721189 (4621, 4622) Novel Protein sim. GBank g1(2137337 pit 46281 - gene Contains protein domain (PF00013) - transcriptfactor mCBP protein - mouse						265018, 264604, 265019, 18108351, 264448,
78415283 (4613, 4614) 87606409 (4615, 4616) Novel Protein sim. GBank gild 78572/regild (4617, 4618) Novel Protein sim. GBank Gractor synthesis 2 Gractor synthesis 3 G						264288, 264766, 264769, 21906765,
### 1983 (4613, 4614) ### 1983 (4613, 4614) ### 1983 (4611, 4614) ### 1983 (4611, 4614) ### 1983 (4611, 4614) ### 1983 (4611, 4619) ### 1983 (4611, 4619) ### 1983 (4611, 4619) ### 1983 (4611, 4629)	_					21906766, 21906767, 21906768, 29148629,
8760408 (4613, 4614) 8760408 (4615, 4616) Novel Protein sim. GBank 978415283 (4613, 4614) 8760408 (4615, 4616) Novel Protein sim. GBank 97857218 (4617, 4618) Novel Protein sim. GBank gi[3878059]emb[CAB17070] - (29942) cDNA EST EMBL:D7344 comes from this gene; cDNA EST EMBL:D7208 comes from this gene; cDNA EST EMBL:D7209 comes from this gene; cDNA EST EMBL:D7209 comes from this gene; cDNA EST EMBL:D7294 comes from this gene; cDNA ENBL:D7294 comes from this gene; cDNA ENBL:D72994 comes from this gene; cDNA ENBL:D72995 comes from th						29148784, 35695917, 265020, 265021,
### 19815283 (4613, 4614) ### 178415283 (4613, 4614) ### 178415283 (4615, 4616) ### 178415283 (4615, 4619) ### 178415283 (4611, 4619) ### 17841688 (4619, 4620) ### 17841688 (4619, 4620) ### 17841688 (4619, 4620) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 17841688 (4621, 4622) ### 178416888 (4621, 4622) ### 178416888 (4621, 4622) ### 178416888 (4621, 4622) ### 1784168888 (4621, 4622) ### 1784168888888888888888888888888888888888						265022, 33657023, 264692, 18108370,
### 1980/1688 4613, 4614 Novel Protein sim. GBank DINCLASSIFIED ### 1980/1688 4617, 4619 Novel Protein sim. GBank ### 1980/1688 4617, 4619 Novel Protein sim. GBank gij3878059 emb CAB17070 - Cobaso somes from this gene; CDNA EST EMBL: D73208 comes from this gene; CDNA EST EMBL: D73209 comes from this gene; CDNA EST EMBL: D73209 comes from this gene; CDNA EST EMBL: D73509 comes from this gene; CDNA EST EMBL: D73509 from this gene; CDNA EST EMBL: D73509 from this gene; CDNA EST EMBL: D7509 from this graph from this gene; CD						18108374, 18108376, 35696423, 35695855,
100 100						264630, 264634, 264635, 264636, 264637,
19415283 (4613, 4614) Wovel Protein sim. GBank gild758722(16(Novel Protein sim. GBank gild758722(16(Novel Protein sim. GBank gild758732(16(Novel Protein sim. GBank gild75873)) WOLLASSIFIED synthase gild758732(16(Novel Protein sim. GBank gild737337(pir)) WOLLASSIFIED synthase gild758732(16(Novel Protein sim. GBank gild737337(pir)) WOLLASSIFIED showel Protein sim. GBank gild737337(pir) Wollamin sprotein domain (PF00013) - transcriptfactor mCBP protein - mouse WOLLASSIFIED WOLLASSIFI						264638, 264639, 18108382, 18108385,
UNCLASSIFIED	1					18108387, 264563, 264566, 264486
87606409 (4615, 4616) Novel Protein sim. GBank g19878059 emb CAB17070 - Goddor synthesis 2 Goddor GBank g13878059 emb CAB17070 - (299942) cDNA EST EMBL:D7208 comes from this gene; cDNA EST EMBL:D7208 comes from this gene; cDNA EST EMBL:D7208 comes from this gene; cDNA EST EMBL:D72094 comes from this gene; cDNA EST EMBL:D72094 comes from this gene; cDNA 23601668 (4619, 4620)		79415283 (4613, 4614)			UNCLASSIFIED	264828
914758732 ref NP_004522.1 pMOCS - molybdenum cofactor synthesis 2 95357218 (4617, 4618) Novel Protein sim. GBank gil3878059 emb CAB17070 - (299942) cDNA EST EMBL.D72906 comes from this gene; cDNA EST EMBL.D7209 comes from this gene; cDNA EST EMBL.D7209 comes from this gene; cDNA EST EMBL.D7209 comes from this gene; cDNA EST EMBL.D72094 comes from this gene; cDNA EST EMBL.D72091 comes from this gene; cDNA EST EMBL.D72094 comes from this gene; cDNA EST EMBL.D72091 comes from this gene; cDNA EST EMBL.D72091 comes from this gene; cDNA EST EMBL.D7209 comes from this gene; cDNA EST EMBL.D72091 comes from this gene; cDNA EST EMBL.D72091 comes from this gene; cDNA EST EMBL.D7209 comes from this gene; cDNA EST EMBL.D72091 comes from this gene; cDNA EST EMBL.D72091 comes from this gene; cDNA EST EMBL.D7209 comes from this gene; cDNA EST EMBL.D72091 comes from this gene; cDNA EST EMBL.D72091 comes from this gene; cDNA EST EMBL.D7209 comes from this gene; cDNA EST EMBL.D72091 comes from this gene; cDNA EST EMBL	_	87606409 (4615, 4616)	Novel Protein sim. GBank		synthase	35696286, 264259, 29331822, 29331824,
Colactor synthesis 2	_		gil4758732 ref NP_004522.1 pMOCS - motybdenum	_		264112, 264512, 264757, 21906754, 264288,
95357218 (4617, 4618) Novel Protein sim. GBank gij3878059jembjCAB17070j - (29942) cDNA EST EMBL:D73444 comes from this gene; cDNA EST EMBL:D73444 comes from this gene; cDNA EST EMBL:D73905 comes from this gene; cDNA EST EMBL:D72944 comes from this gene; cDN (29942) cDNA EST EMBL:D73942 comes from this gene; cDNA EST EMBL:D72944 comes from this gene; cDN (29601668 (4619, 4620)			cofactor synthesis 2			264690, 27486264, 264631, 264634, 264404
(299942) cDNA EST EMBL:D73444 comes from this gene; cDNA EST EMBL:D73456 comes from this gene; cDNA EST EMBL:D720905 comes from this gene; cDNA EST EMBL:D7204 comes from this gene; cDNA EST EMBL:D72044 comes from this gene; cDN [28601668 [4619, 4620]		95357218 (4617, 4618)	Novel Protein sim. GBank gij3878059jembjCAB17070j -	3	glycoprotein	18108397, 22278996, 22278997, 22278998,
CDNA EST EMBL: D70906 comes from this gene; cDNA EST EMBL: D7208 comes from this gene; cDNA EST EMBL: D7208 comes from this gene; cDNA EST EMBL: D72084 comes from this gene; cDN 29601668 (4619, 4620) Novel Protein sim. GBank gi[2137337 pir 148281 - gene Contains protein domain (PF00013) - transcriptfactor mCBP protein - mouse KH domain			(299942) cDNA EST EMBL:D73444 comes from this gene;			22278999, 60432049, 29331822, 29331826,
EST EMBL:D7208 comes from this gene; cDNA EST EMBL:D72084 comes from this gene; cDN EMBL:D72944 comes from this gene; cDN EMBL:D72944 comes from this gene; cDN B7721189 (4619, 4620) MCBP protein sim. GBank gi[2137337]pir 148281 - gene Contains protein domain (PF00013) - transcriptfactor MCBP protein - mouse KH domain			cDNA EST EMBL:D70905 comes from this gene; cDNA	_		60432289, 66712502, 60432229, 60433356,
EMBL:D75030 comes from this gene; cDNA EMBL:D72944 comes from this gene; cDN 79601668 (4619, 4620) 87721189 (4621, 4622) MCBP protein sim. GBank gi[2137337]pir 148281 - gene Contains protein domain (PF00013) - transcriptfactor MCBP protein - mouse KH domain			EST EMBL:D72208 comes from this gene; cDNA EST			60433438, 65274444, 265010, 264600,
EMBL:D72944 comes from this gene; cDN 79601668 (4619, 4620) 87721189 (4621, 4622) Novel Protein sim. GBank gi[2137337]pir 148281 - gene Contains protein domain (PF00013) - transcriptfactor MCBP protein - mouse KH domain			EMBL:D75030 comes from this gene; cDNA EST			264681, 264448, 264683, 264288, 21906766,
79601668 (4621, 4620) 87721189 (4621, 4620) MOVEL Protein sim. GBank gi[2137337]pir 148281 - gene Contains protein domain (PF00013) - transcriptfactor KH domain			EMBL:D72944 comes from this gene; cDN			21906768, 265020, 264691, 264692, 264693,
UNCLASSIFIED 87721189 (4621, 4622) Novel Protein sim. GBank gi[2137337 pir 148281 - gene Contains protein domain (PF00013) - transcriptfactor mCBP protein - mouse KH domain	1	70004 0404) 00040007				65274620, 65274791
B7721189 (4621, 4622) Novel Protein sim. GBank gi[2137337 pir 148281 - gene Contains protein domain (PF00013) - transcriptfactor KH domain	1	(300 1000 (4013, 40ZU)			UNCLASSIFIED	264508
KH domain	_	187721189 (4621, 4622)	Novel Protein sim. GBank gi[2137337]pir[]148281 - gene	Contains protein domain (PF00013) - t	transcriptfactor	18108397, 56182575, 22278996, 56994075,
264907, 56182435, 264510, 264511, 265008, 264510, 265011, 265008, 265011, 265008, 265011, 265009, 60433431, 264511, 265011, 264611, 26403, 265011, 26403, 265011, 26403, 265011, 26403, 265011, 26403, 265011, 26403, 26436, 26436, 26436, 26436, 264631, 264638, 264688, 2646			mCBP protein - mouse	KH domain		264259, 29331824, 29331827, 264508,
264512, 265007, 265009, 60433431 33109954, 265010, 265011, 264603, 265011 18108351, 264762, 254663, 264288, 264368, 264688, 33657023, 20281149, 202811069, 264628, 264628, 264632, 264632, 264632, 264632, 264638, 264688, 264638, 264688, 264						264907, 56182435, 264510, 264511, 265006,
33109954, 265010, 265011, 264603, 265011, 264603, 265011, 264603, 265011, 264603, 265011, 264603, 265011, 264603, 264361, 264681, 26481	_					264512, 265007, 265008, 265009, 60433438,
18108351, 264762, 2642683, 264366 264686, 33657023, 20281149, 20281069, 264628, 263972, 55811576, 35696423, 20281071, 264532, 264636, 18108385, 18108387, 87168518, 22279000, 264563,						33109954, 265010, 265011, 264603, 265017,
264686, 33657023, 20281149, 20281069, 264628, 263972, 55811576, 35696423, 20281071, 264532, 254536, 18108385, 18108387, 87168518, 22279000, 264563, 2645645	_					18108351, 264762, 264683, 264288, 264369,
264628, 263972, 55811576, 35696423, 20281071, 264532, 284636, 18108385, 18108387, 87188518, 22279000, 264563,	_					264686, 33657023, 20281149, 20281069,
20281071, 264632, 264636, 18108385, 18108387, 87168518, 22279000, 264563,	_					264628, 263972, 55811576, 35696423,
18108387, 87158518, 22279000, 284563,						20281071, 264632, 264636, 18108385,
						18108387, 87168518, 22279000, 264563,

90042533 (4625, 4626) Novel Protein sim. CBank gij3043626jdbjjBAA254771 - 90042533 (4625, 4626) Novel Protein sim. CBank gij3043626jdbjjBAA254771 - 94313401 (4622, 4636) Novel Protein sim. CBank gij304362jdbjjBAA254771 - 94313401 (4622, 4630) Novel Protein sim. CBank gij50571607111 - 94312191 (4629, 4630) 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD44488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein Sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein Sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein Sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein Sim. CBank gij5531827[gb]AAD4488.11 - 94312191 (4631, 4632) Novel Protein	2242	1075 40501 145301 45341	1013000 47 10011001 - 100 - 1 - 10110 11111 11111 11111 111111 111111				Γ
94313401 (4627, 4526) Novel Protein sim. GBank gij3043628(db)[BAA25477] - 94313401 (4627, 4528) (A8011123) KIAA0351 protein (Horms suptered) (AL055399) (47796117.2 (CGL-51) (Hormo saplens) (AL055399) (47796117.2 (CGL-51) (Hormo saplens) (AF078856) Protein sim. GBank gil5531827(gb)AAD44488 11 - UEX dominin (PF00789) - ghccpprotein (AF078856) p47 (Hormo saplens)	4	(+50+,520+) langer	Unknown gene product [Homo sapiens]			301823/3, 309940/3, 33096286, 222/8996, 22278997, 22278999, 264259, 29331822	
80042533 (4625, 4626) Novel Protein sim. GBank gj30438298pjBAA224771 - 94313401 (4627, 4628) Novel Protein sim. GBank gj30438298pjBAA224771 - (AL055298) (47780172 (CGI-51) [Homo saplens] 94312191 (4629, 4630) 94312191 (4621, 4632) Novel Protein sim. GBank gj3531927jgbJAD34488 11 - UBX domain UBX domain						29331824, 29331825, 29331826, 29331827,	
80042333 (4622, 4628) Novel Protein sim. GBank gil3043929(da)[BAA25477] - 94313401 (4627, 4628) Novel Protein sim. GBank gil3043929(da)[BAA25477] - (AL053398) J796117.2 (CG1-51) [Homo saplens] 94312191 (4631, 4632) Novel Protein sim. GBank gil5531627[gb]AAD4488 1] - (AL05398) J79617.2 (CG1-51) [Homo saplens] (AF0778656) p47 [Homo saplens] (AF0778656) p47 [Homo saplens] (AF0778656) p47 [Homo saplens]						52644045, 60432229, 60433356, 55812038.	
80042333 (4625, 4626) Novel Protein sin. CBank gij304352gidbijBAA254771 - (AL035389) dJ796117 2 (CGI-51) [Homo sapiens] 80430119 (4629, 4630) 94312191 (4629, 4630) 94312191 (4631, 4632) Novel Protein sin. CBank gij5361927[gbjAAD44468 1] - (Contains protein domain (PF00789) - glycoprotein (AF078859) p47 [Homo sapiens]						33109954, 21906754, 87168474, 265018,	
80042533 (4625, 4626) Novel Protein sin. GBank gij3043628[dob]gAA25477 - 94313401 (4627, 4628) Novel Protein sin. GBank gij3043628[dob]gAA25477 - (AL035398) dJ796117 2 (CG1-51) [Homo saplens] 80430119 (4629, 4630) 94312191 (4629, 4630) (AL035398) p47 [Homo saplens] (AF078856) p47 [Homo saplens] (AF078856) p47 [Homo saplens]						18108351, 264288, 52644229, 21906765,	
94313401 (4627, 4626) Nover Protein sim. GBark gij304626[dbjjBAA25477] - (AL035398) dJ796117.2 (CGL-S1) [Homo saplens] (AL035398) dJ796117.2 (CGL-S1) [Homo saplens] (AL035398) dJ796117.2 (CGL-S1) [Homo saplens] (AF073856) p47 [Homo saplens] (AF073856) p47 [Homo saplens] (AF073856) p47 [Homo saplens]						21906766, 21906767, 21906768, 21906769,	
80042533 (4625, 4626) Novel Protein sim. GBank gij3043626(db)j8A4254711- [AL035398) dJ798117.2 (CGI-51) [Homo sapients] 94312191 (4631, 4632) Novel Protein sim. GBank gij5531827[gb]AAD44488 11- (AL035398) p47 [Homo sapients] UNCLASSIFIED 94312191 (4631, 4632) Novel Protein sim. GBank gij5531827[gb]AAD44488 11- UBX domain (PF00789) - glycoprotein UBX domain						265022, 60170615, 33657023, 27486261,	
80430119 (4621, 4630) Novel Protein sim. GBank gij3043628(db) BAA2477 - (AL035398) dJ7961172 (CGI-51) [Homo sapiens] (AL035398) dJ796172 (CGI-51) [Homo sapiens]						27486264, 35696423, 35695855, 18108385,	
94313401 (4627, 4628) (World Trotein Sin, Galank gijs531827) (AL035398) dJ796172 (CGI-51) [Homo sapiens] 94313401 (4627, 4628) (World Protein sin, GBank gijs53714[mb]C/AB51401.11- (AL035398) dJ796172 (CGI-51) [Homo sapiens] 94312191 (4631, 4632) Nove Protein sin, GBank gijs531827] (gb)AAD44488.1 - (AF078856) p47 [Homo sapiens] UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED (AF078856) p47 [Homo sapiens]	2242	90043633 4636	The same of the sa			22279000, 22279002	Ţ
94313401 (4627, 4629) Novel Protein sim. CBank gij559714pm0 sapiens] 80430119 (4629, 4630) 94312191 (4631, 4632) (AF078856) p47 [Homo sapiens] UBX domain UBX domain UBX domain	2 27		Novel Protein sim. GBank gij3043626jdbj BAA25477 - (AB011123) KIAA0551 protein [Homo sapiens]			263981	
(AL035398) dJ796i17.2 (CGI-51) [Homo sapiens] 80430119 (4629, 4630) 64512191 (4631, 4632) Novel Protein sim. GBank gil5531827[gb AAD4488.1] - Contains protein domain (PF00789) - glycoprotein (AF078856) p47 [Homo sapiens] LBX domain	2314		Novel Protein sim. GBank gil5596714 emb CAB51401.1 -	UNCL	LASSIFIED	52644507, 52646365, 52646842, 22278996,	Π
80430119 (4629, 4630) 94312191 (4621, 4632) Novel Protein sim. GBank gij5531627lgbl,AAD4488.1 - Contains protein domain (PF00789) - glycoprotein (AF078856) p47 [Homo sapiens]			(AL035398) dJ796117.2 (CGI-51) [Homo sapiens]			22278997, 22278999, 264259, 52645080.	
80430119 (4629, 4630) 94312191 (4631, 4632) Novel Protein sim. GBank gijs531827jgb AAD44488.1 - Contains protein domain (PF00789) - glycoprotein (AF078856) 947 [Homo sapiens] UBX domain						29331822, 29331824, 60432289, 29331826,	
80430119 (4629, 4630) 94312191 (4631, 4632) Novel Protein sim. GBank gijs531827igb AAD44488.1 - (AF078856) p47 [Homo septens] UBX domain UBX domain		-				29331827, 29331828, 35696052, 33656970,	
80430119 (4629, 4630) 94312191 (4631, 4632) Novel Protein sim, GBank gil5531827(gbl/AAD4488.1 - Contains protein domain (PF00789) - glycoprotein (AF078856) p47 [Homo sapiens]						52644045, 265008, 264593, 60433356,	
80430119 (4629, 4630) 94312191 (4631, 4632) Novel Protein sim. GBank gil5531827[gb]AAD44488.1 - Contains protein domain (PF00789) - glycoprotein UBX domain						60433438, 264758, 33109954, 265010,	
80430119 (4629, 4630) 94312191 (4631, 4632) Novel Protein sim. GBank gi 5531827 gb AAD44488.1 - Contains protein domain (PF00789) - glycoprotein (AF078856) p47 [Homo sapiens]						265017, 265018, 265019, 264288, 264369,	
80430119 (4629, 4630) 94312191 (4631, 4632) Novel Protein sim. GBank gil5531827[gb AAD44488.1] - Contains protein domain (PF00789) - glycoprotein (AF078856) p47 [Homo sapiens]	_					21906765, 21906766, 21906768, 35695917,	
80430119 (4629, 4630) UNCLASSIFIED 94312191 (4631, 4632) Novel Protein sim. GBank gil5531827[gb AAD4488.1] - Contains protein domain (PF00789) - glycoprotein UBX domain						52644150, 33657023, 33657109, 52645129,	
80430119 (4629, 4630) 94312191 (4631, 4632) Novel Protein sim. GBank gil5531827]gb[AAD44488.1]- (AF078856) p47 [Homo sapiens] Contains protein domain (PF00789) - glycoprotein UBX domain						33657349, 35695763, 18108374, 35696423,	
94312191 (4631, 4630) UNCLASSIFIED 94312191 (4631, 4632) Novel Protein sim. GBank gil5531827]gb AAD4488.1 - UBX domain (AF078856) p47 [Homo sapiens] UBX domain						35695855, 52644332, 22279000, 22279002,	
80430119 (4629, 4630) 94312191 (4631, 4632) Novel Protein sim. GBank gil5531827 gb AAD4488.1 - Contains protein domain (PF00789) - glycoprotlein (AF078856) p47 [Homo sapiens]						264563, 264567	
64312191 (4631, 4632) Novel Protein sim. GBank gil5531827]gb AAD4488.1 - Contains protein domain (PF00789) - glycoprotein (AF078856) p47 [Homo sapiens]	2315	80430119 (4629, 4630)		UNCL	LASSIFIED	264905, 264906, 264767, 264768, 264693,	П
94312191 (4631, 4632) Novel Protein sim. GBank gil5531827[gb]AAD44488.1 - Contains protein domain (PF00789) - glycoprotein (AF078856) p47 [Homo sapiens] UBX domain		_				55811576, 264635, 56182323, 18108385	
UBX domain	2316			Contains protein domain (PF00789) - glycop	protein	52644507, 52645156, 52646365, 22278994,	П
22278997, 22278999, 2 29331822, 2654060, 29331824, 2 29331822, 2654060, 29331827, 3596052, 2 29331822, 2564060, 26331827, 3596052, 2 264906, 264907, 29331830, 25644 56182435, 264511, 265007, 26500 60170831, 6043348, 21906754, 5 3109954, 33657084, 264683, 2644296, 8 2564429, 21906764, 26468, 264617, 26510 264681, 264763, 264429, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 21906764, 264689, 265017, 2650170				UBX domain		22278995, 22278996, 56994075, 35696286,	
29331822, 52645080, 29331824, 2 29331827, 35696052, 2 264906, 264907, 26331830, 52644 56182435, 264511, 265007, 25500 60170831, 60433438, 2160674, 25500 60170831, 60433438, 2160674, 25601 5264010, 81166559, 256017, 26501 5264010, 81166559, 256017, 26501 5264010, 81166559, 21606764, 264689, 219 52644229, 21906764, 264689, 219 52644229, 21906764, 264689, 219 53657023, 18108362, 52645129, 33657049, 35657631, 18108385, 264567						22278997, 22278998, 22278999, 264259,	
29331826, 29331827, 35696052, 2 264906, 264007, 29331830, 25644 56182435, 284511, 2565007, 26500 60170831, 60433438, 21906754, 265017, 26501 265010, 87168559, 265017, 26501 264681, 264429, 21906764, 264683, 2564429, 21906764, 264681, 21906764, 264681, 21906764, 264681, 21906764, 264681, 21906764, 264683, 21906764, 264687, 2646854, 264687, 2646854, 264687, 2646						29331822, 52645080, 29331824, 29331825,	
264906, 264907, 29331830, 52644 56182435, 26451, 265007, 26500 6017081, 1265007, 26500 6017081, 1265007, 26500 6017081, 1265007, 26500 6017081, 1265007, 26500 6017081, 264029, 21906764, 264029, 8 6264010, 81168559, 265017, 26501 626483, 26448, 26683, 6264229, 21906764, 26468, 219 6264681, 26473, 18108529, 21906768, 2906768, 2906768, 2906769, 2						29331826, 29331827, 35696052, 29331828,	
\$6182435, 264511, 265007, 26500 60170831, 6043248, 21906764, 55007, 26501 2010954, 33057094, 52644296, 8 23109954, 33057094, 265017, 26501 26461, 264763, 264448, 264630, 21906764, 264689, 219 21906766, 21906767, 21906767, 21906767, 21906767, 21906767, 21906767, 21906768, 2 21906766, 21906767, 21906767, 21906768, 2 21906766, 21906767, 21906768, 2 21906766, 21906767, 21906768, 2 21906766, 21906767, 21906768, 2 21906766, 21906767, 21906768, 2 21906766, 21906767, 21906768, 2 21906768, 21906768, 21906768, 2 21906768, 21906768, 2 21906768, 21906768, 2 21906768, 21906768, 2 21906768, 21906768, 2 21906768, 21906768, 2 21906768, 21906768, 2 2279000, 2 2275000, 2 20106768, 2 2010						264906, 264907, 29331830, 52644045,	
60170831, 60433438, 21906764, 5 3109954, 33557084, 52644296, 8 265010, 8716859, 265017, 26501 264681, 264763, 26448, 265017, 26501 264681, 264763, 26448, 265021, 2650						56182435, 264511, 265007, 265008, 265009,	-
33109954, 33657084, 52644296, 8 265010, 87166559, 265017, 265017 264681, 264763, 26448, 264683, 265017, 265017 264681, 264763, 26448, 264683, 21906766, 21906766, 21906766, 21906768, 23695917, 265020, 265021, 52644 33657023, 18108362, 265021, 52644 33657349, 35695763, 18108370, 18108370, 18108385, 18108373, 18108385, 18108385, 1810831, 22279000, 265057						60170831, 60433438, 21906754, 52646317,	
265010, 87166559, 265017, 26501 264681, 284763, 284448, 284683, 2844229, 21906764, 284689, 219 21906766, 21906768, 2 35695917, 265020, 265021, 52644 33657349, 35695763, 18108370, 18108370, 18108385, 1810855, 1810						33109954, 33657084, 52644296, 87168474,	
264681 284763 284448 294683. 2644229, 21906764, 284689, 219 21906766, 21906767, 21906768, 2 3569517, 265021, 25644 33657349, 35695763, 18108362, 18108370, 18108385, 1810836, 18108385, 1810836, 18108385, 1810836, 1810838						265010, 87168559, 265017, 265018, 265019,	
5264429, 21906764, 26468, 219 21906766, 21906767, 21906768, 219 21906766, 21906767, 21906768, 219 31657023, 18108362, 52645129, 3 31657349, 3569563, 18108370, 1 3569423, 3569585, 284611, 284 52644332, 83373044, 18108385, 1 87168518, 60432113, 22279000, 2				-		264681, 264763, 264448, 264683, 264369,	
21906766, 21906767, 21906768, 2 35695917, 265021, 2564129, 3 33657349, 35695763, 18108370, 1 35694332, 3569585, 26431, 264 52644332, 3373044, 18108385, 1 87168518, 60432113, 22279000, 2						52644229, 21906764, 264689, 21906765,	
35695917, 265020, 265021, 52644 33657023, 18108362, 52645129, 3 33657349, 35695763, 18108370, 1 35696423, 35695855, 264631, 264 52644332, 83373044, 18108385, 1 87168518, 60432113, 22279000, 2						21906766, 21906767, 21906768, 21906769,	
33657023, 18108362, 52645129, 3 33657349, 35695763, 18108370, 1 35696423, 35695855, 264631, 264 52644332, 83373044, 18108385, 1 87168518, 60432113, 22279000, 2	_					35695917, 265020, 265021, 52644150,	
33657349, 35695763, 18108370, 1 35696423, 35695855, 264631, 264 52644332, 83373044, 18108385, 1 87168518, 60432113, 22279000, 2						33657023, 18108362, 52645129, 33657182,	
35696423, 35695855, 264631, 264 52644332, 83373044, 18108385, 1 87168518, 60432113, 22279000, 2 264567						33657349, 35695763, 18108370, 18108376,	
52644332, 83373044, 18108385, 1 87168518, 60432113, 22279000, 2 264567						35696423, 35695855, 264631, 264556,	
87168518, 60432113, 22279000, 2 264567						52644332, 83373044, 18108385, 18108387,	
264567				-		87168518, 60432113, 22279000, 264566,	
						264567	\neg

	_								_				1																T	٦
22278998, 60432049, 264910, 60432229, 264686, 264687, 264688, 264689, 264558, 18108385	265006, 264910	264488, 264569, 18108396, 52646365, 22278994, 22278996, 22278996, 56994075, 35696286, 22278997, 22278998, 264259,	52645080, 29331825, 29331826, 29331827, 29331828, 29331830, 56182435, 60170831	60432229, 60431735, 33657402, 21906754,	52644296, 87168474, 265011, 87168559,	[255017, 265018, 265019, 18108351, 254448, 18108354, 264288, 264369, 52644229.	21906764, 21906765, 21906766, 21906767.	21906768, 21906769, 265021, 265022,	52644150, 33657023, 52645129, 33657109,	27486264, 33657349, 35695763, 18108370,	18108376, 18108379, 35696423, 264558,	2337,3044,	22278994, 60432049, 60432289, 29331827,	264511, 265008, 52646317, 265017, 265019,	21906765, 18108372, 18108387, 22279002	264488, 264687, 18108394, 264689,	21906765, 18108397, 18108398, 21906767,	21906768, 65274791, 22278995, 35695855,	22278998, 265021, 265022, 264510, 265006,	254511, 264512, 265008, 60170515, 264555, 264556, 264666 264666 264666	60432229 33657023 264557 264558	264693, 60433356, 264559, 60433438,	29331824, 18108365, 18108348, 18108384,	29331825, 18108385, 33109954, 29331827,	56526486, 29146499, 265011, 60432113,	265017, 265018, 264508, 264563, 264482,	264509, 18108351, 264448, 264907, 264682,	18108370, 264683, 264908, 264288, 264909, 18108384, 264486, 264567	22278996, 264907, 264511, 264757, 18108351, 264768, 264638	
UNCLASSIFIED	UNCLASSIFIED												kinase			UNCLASSIFIED		_											dehydrogenase	
																													Contains protein domain (PF00106) - dehydrogenase short chain dehydrogenase	
		2319 95101781 (4637, 4638) Novel Protein sim. GBank gi 5262613 emb CAB45746.1 - (AL080155) hypothetical protein [Homo sapiens]											2320 91622426 (4639, 4640) Novel Protein sim. GBank	gi/728837/sp/P39194/ALU7_HUMAN - !!!! ALU SUBFAMILY	SQ WARNING ENTRY !!!!	Novel Protein sim. GBank gij3873837 jemb CAB02700 j -	(Z81029) Similarity to S.pombe hypothetical protein	C1D4.09C (SW:Q10154); cDNA EST EMBL:100543 comes	good only Est Ewel 194234 and the this	gene, Colve Est Embl. 101321 comes nom uns gene; CONA EST EMBL-10238 som									87803165 (4643, 4644) Novel Protein sim. GBank gi 5678957 emb CAB51685.1 - (AL109630) BACR7A4.y [Drosophila melanogaster]	
2317 87020571 (4633, 4634)	79959879 (4635, 4636)	95101781 (4637, 4638) 1											91622426 (4639, 4640)			94320377 (4641, 4642)													87803165 (4643, 4644)	
2317	2318	2318		_									2320		- 1	2321													2322	

22278994, 22278995, 22278997, 60432049, 264259, 29331822, 33656970, 224509, 56182435, 264511, 265008, 60433356, 60433438, 55812038, 33109954, 21906754, 265017, 265017, 265019, 264760, 264681, 18108351, 265075, 256097, 21906767, 21906768, 55811957, 35695917, 265021, 33657023, 18108362, 27486262, 55811576, 26431, 264555, 83373044, 8716818, 60432113, 22279002	264592, 264593, 265020		265006, 264759, 35695855, 56182323		264259, 264508, 264905, 264906, 264907, 264908, 265007, 264512, 264910, 264758, 265010, 264766, 264768, 264769, 33657023,	264693, 264628, 264631, 264634, 264638, 264639, 264486	60433438, 264595, 265017, 264766, 264692,	264629, 264635, 264636, 264638, 56182323, 60432113, 264566	265017, 264685, 60432113, 264088	265009		265006, 265008, 265018, 264448, 264764, 21906765, 35695917, 35695855, 264636, 22279000, 264566	56182575, 56994075, 29331826, 29331828,	264107, 33637402, 07165333, 254650. 35695917, 265021, 33657023, 263976	56182575, 29331825, 21906768, 264636, 83373044
- Lde		ATPase_associated	cyto450	UNCLASSIFIED	UNCLASSIFIED		UNCLASSIFIED		UNCLASSIFIED	transport	ATPase_associated		UNCLASSIFIED		- dehydrogenase
Contains protein domain (PF00226) - eph DnaJ domain			Contains protein domain (PF00067) - cyto450 Cytochrome P450				Contains protein domain (PF00735) - UNCLASSIFIED	Cell division protein							Contains protein domain (PF00441) - dehydrogenase Acyl-CoA dehydrogenase
.1 IN CHROMOSOME II		Bank gi[5419865 emb CAB46377.1 - tical protein [Homo sapiens]	20	88081648 (4653, 4654) Novel Protein sim. GBank gil4240227 dbj BAA74892.1 - (AB020676) KIAA0869 protein [Homo sapiens]	83388428 (4655, 4656) Novel Protein sim. GBank gi[1245105 (U46463) - glutamine repeat protein-1 [Mus musculus]			6_MOUSE - DIFF6 PROTEIN		86990463 (4661, 4662) Novel Protein sim. GBank gilj679136[gb AAD46874.1 AF16093 - (AF160934) RcDNA I D14189 (Drosophia melanonaster)	87784182 (4663, 4664) Novel Protein sim. GBank gil 2104452 embl CAB08779 -	(Z95397) unknown [Schizosaccharomyces pombe]	88206958 (4665, 4666) Novel Protein sim. GBank gil3879985 emb CAA92691.1 -	(Z68318) cDNA EST CEMSD62F comes from this gene; cDNA EST EMBL:C07930 comes from this gene; cDNA EST EMBL:C09493 comes from this gene; cDNA EST yk415e8.3 comes from this gene; cDNA EST yk415e8.5 comes from this gene; cDNA EST yk415e8.55	94319788 (4667, 4668) Novel Protein sim. GBank gil4966270[gb]AAB52261.2 - (U97002) similar to acyl-CoA dehydrogenases and epoxide hydrolases; Pfam domain PF00441 (Acyl-CoA_dh). Score=57.4, E-value=1.7e-16, N=2; contains similarity to Pfam domain PF00702 (Hydrolase), Score=57.4, E-
94840445 (4645, 4646)	86633607 (4647, 4648)	88165074 (4649, 4650)	84390962 (4651, 4652)				87604478 (4657, 4658)		87335396 (4659, 4660)	2331 86990463 (4661, 4662)					
2323	2324	2325	2326	2327	2328		2329		2330	2331	2332		2333		2334

2335	80046103 (4669, 4670) Novel Protein sim. Gladulin-binding or	Novel Protein sim. GBank gij3283350 (AF062378) - calmodulin-binding protein SHA1 (Mus. musculus)	Contains protein domain (PF00612) - struct	struct	18108351, 21906769, 264555	_
2336		95196121 (4671, 4672) Novel Protein sim. GBank gi[1929056 emb CAA72805 - (Y12090) putative 3.4-dihydroxy-2-butanone kinase [Lycopersicon esculentum]		kinase	264907, 35695917, 18108379	
2337	95345810 (4673, 4674)	95345810 (4673, 4674) Novel Protein sim. GBank gil4495063 emb CAB39181.1 - (725986) dJ108K11.3 (similar to yeast suppressor protein SRP40) [Homo sapiens]		UNCLASSIFIED	35696286, 22278999, 56182181, 29331825, 60424269, 56182435, 33657402, 55812038, 55811386, 265017, 265018, 265019, 21906768, 35695917, 264691, 33657103, 33657109, 263972, 35696423, 35695855, 60432119,	
2338	87634045 (4675, 4676)	87634045 (4675, 4676) Novel Protein sim. GBank gil2224689 db BAA20829 - (AB002372) KIAA0374 [Homo sapiens]	Contains protein domain (PF00323) - UNCLASSIFIED Marnmalian defensin		18108394, 29331822, 66714117, 60432289, 264906, 56182435, 265009, 60433438, 264596, 265010, 265019, 18108354, 264288, 264369, 55811957, 265021, 33657023, 263976, 55811576, 264632, 56182323, 264639	
2339	85663319 (4677, 4678)	85663319 (4677, 4678) Novel Protein sim. GBank gij3873550 emb CAA22127 - (AL033534) serine-rich protein [Schizosaccharomyces pombe]		UNCLASSIFIED	35696286, 264592, 264369, 264691, 264558	
2340	90937716 (4679, 4680)				65274572, 22278994, 36696286, 22278997, 22278999, 264259, 29331822, 60432289, 29331892, 60432289, 29331896, 263019, 265009, 33657402, 33109954, 265017, 265018, 264786, 264695, 3696432, 87168518, 22279000	
2341					264259, 264908, 264909, 264682, 22279000	_
		95334968 (4683, 4684) Novel Protein sim. GBank gij3874563jemb[CAB02797] - (Z81042) similar to Yeast hypothetical protein YEY6 like; cDNA EST yk206h5.3 comes from this gene; cDNA EST yk206h5.5 comes from this gene; cDNA EST yk303h1.3 comes from this gene; cDNA EST yk303h1.5 comes from this gene; cDNA EST yk303h1.5 comes from this gene; cDNA	Contains protein domain (PF00400) - kinase WD domain, G-beta repeat		264488, 65274572, 22278995, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 66714117, 29331825, 60432289, 35696052, 264905, 264509, 29331830, 265005, 264905, 264509, 29331830, 265006, 265017, 265019, 264448, 264369, 265017, 265019, 264448, 264369, 265917, 265020, 265022, 5264150, 33657023, 65274620, 33657109, 18108370, 18108376, 264556, 264568, 264564, 264566, 264566, 264567, 264568, 264566, 264566, 264568, 264688, 264688, 264688, 264688, 264688, 264688, 264688, 264688, 264684, 264684, 264684, 264684, 264684, 264684, 264684, 264684, 264684, 264	, <u></u>
2343	87775448 (4685, 4686)	2343 87775448 (4685, 4686) Novel Protein sim. GBank gil4929741[gb AAD34131.1 AF15189 - (AF151894) CGI-136 protein [Homo sapiens]		UNCLASSIFIED	264907, 264512, 265011, 264683	
2344				UNCLASSIFIED	264758	
2345		94319799 (4689, 4690) Novel Protein sim. GBank gil2506307[sp[P13944 CA1C_CHICK - COLLAGEN ALPHA von Willebrand factor type A domain [1(XII) CHAIN PRECURSOR (FIBROCHIMERIN)		collagen	264488, 264259, 66712502, 264759, 83373044, 264566	

2346 34131220 (4691, 4692) Nove Protein sim. GBank gil 192906 [emb] CAA72805] Caerins protein domain (PF0315) - protessenth) 2340 36136120 (4695, 4699) Screvisiae glucose repression mediator protein 2341 36530367 (4695, 4699) Screvisiae glucose repression mediator protein 2342 36530367 (4695, 4699) Novel Protein sim. GBank gil 192906 [emb] CAA72805] Caerina protein domain (PF0315) Protein sim. GBank gil 192906 [emb] CAA72805] Caerina protein sim. GBank gil 192906 [emb] CAA72841] Caerina protein sim. GBank gil 192906 [emb] CAA72841] Caerina sim. GBank gil 192906 [emb] CAA72841] Caerina sim. GBank gil 192906 [emb] CAA7287 Caerina sim. GBank gil 192906 [emb] CAA7284 Caerina sim. GBank gil 192906 [emb] CAA7284	35696286, 22278998, 264259, 35696052. 28331828, 33657402, 60433356, 33109954, 87168559, 264603, 265019, 18108351, 264681, 264685, 21906768, 265021, 33657109, 55811576, 35695855, 264637, 52644332, 264557, 83373044, 22279000,	22278997, 264511, 264683, 264684, 264768, 264687, 264688, 264691, 264692, 55811576	18108394, 35696286, 264259, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 264512, 265008, 265009, 264910, 264591, 264592, 264593, 264593, 264591, 264691, 264762, 18108351, 264764, 264288, 264768, 264591, 264693, 264693, 264638, 3659817, 264693, 264631, 264631, 264636, 264636, 264637, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264681, 264486	35696052, 29146499, 264909, 264369	22276998, 22278999, 264259, 29331822, 29331824, 29331825, 29331827, 29331828, 33109954, 21906754, 265010, 87168559, 265018, 265019, 264761, 264681, 264691, 264692, 35695855, 87168518, 22279000, 22279002, 264482	56182575, 264909, 265006, 264558	264259, 264448	29331826, 55812038, 265019, 264692, 264636	29331824, 264908, 265006, 265008
4689, 4700] Novel Protein sim. GBank gil1255411 (U33153) - one short Contains protein domain (PF00515) - freeploted of weak similarity (15. cerevisiae proteises A inhibitor) TPR Domain (17.12080) putates stillarity (18.5. deceleration of weak similarity (18.5. deceleration of weak similarity (18.5. deceleration of weak similarity (18.5. deceleration) and another short region of weak similarity (18.5. deceleration) and another short region of weak similarity (18.5. deceleration) and another short region of the region region of the short region of the shor	proteaseinhib		kinase			kinase	UNCLASSIFIED	dna_ma_bind	UNCLASSIFIED
4691, 4692) Novel Protein sim. GBank gil1255411 (U53153) - one short region of weak similarity to S. cerevisiae professe A inhibitor (SP:P14922) (Caenonabditis elegans) 4693, 4694) 4695, 4696) Novel Protein sim. GBank gil1929056[emb[CAA72805] - (Y12090) putative 3,4-dihydroxy-2-butanone kinase (Lycopersion esculentum) (AL050062) hypothetical protein [Homo sapiens] (AL050062) hypothetical protein [Homo sapiens] 4699, 4700) ABB WARNING ENTRY !!!! (4701, 4702) Novel Protein sim. GBank gil4884106[emb]CAB43254.1] - (AL050062) hypothetical protein [Homo sapiens] SB WARNING ENTRY !!!! (4703, 4704) Novel Protein sim. GBank gil488106[emb]CAB43254.1] - (BB WARNING ENTRY !!!! (4703, 4704) Novel Protein sim. GBank gil4889] SB WARNING ENTRY !!!! (4705, 4706) Novel Protein sim. GBank gil4889] SB WARNING ENTRY !!!! (4705, 4706) Novel Protein sim. GBank gil4889] SB WARNING ENTRY !!! (4705, 4706) Novel Protein sim. GBank gil4889] (4707, 4706) Novel Protein sim. GBank gil4889] (4707, 4706) Novel Protein sim. GBank gil4393870] (4707, 4706) Novel Protein sim. GBank gil731637]	Contains protein domain (PF00515) - TPR Domain						Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)
	1891, 4692) Novel Protein sim. GBank gil1255411 (U53153) - one short region of weak similarity to S. cerevisiae protease A inhibitor 3 (SP:P01094) and another short region of weak similarity to S. cerevisiae glucose repression mediator protein (SP:P14922) [Caenorhabditis elegans]	1693, 4694)	f695, 4696) Novel Protein sim. GBank gil 1929056 emb CAA72805 - (Y12090) putative 3,4-dihydroxy-2-butanone kinase [Lycopersicon esculentum]	4697, 4698) Novel Protein sim. GBank gil4884106 emb CAB43254.1 - (AL050062) hypothetical protein [Homo sapiens]	1 699, 4700)	4701, 4702) Novel Protein sim. GBank gij728832jspjP39189JALU2_HUMAN - !!!! ALU SUBFAMILY SB WARNING ENTRY !!!!	4703, 4704) Novel Protein sim. GBank gij731637[splP38760 YHH5_YEAST - HYPOTHETICAL 75.9 KD PROTEIN IN SPO13-ARG4 INTERGENIC REGION	4705, 4706) Novel Protein sim. GBank gil1346955 sp P48809 RB27_DROME - HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN 27C (HNRNP 48) (HRP48.1)	4707, 4708) Novel Protein sim. GBank gij731637 sp P38760 YHH5_YEAST - HYPOTHETICAL 75.9 KD PROTEIN IN SPO13-ARG4 INTERGENIC REGION

(4711, 4712 <u>)</u>	2355 91638786 (4709, 4710) Novel Protein sim. GBank gil4938503 emb CAB43861.1 - (AL078465) hnRNP-like protein [Arabidopsis thaliana] 2356 95327688 (4711, 4712) Novel Protein sim. GBank gil5138920 gb AAD40377.1 - (AF092135) PTD014 [Homo sapiens]	Contains protein domain (PF00076) - dna_rna_bind RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	56994075, 22278996, 35696286, 22278999, 264259, 29331826, 29331828, 29331828, 29331828, 29331828, 2931828, 2931828, 2931828, 293146498, 264905, 264908, 265001, 264687, 21906765, 21906767, 21906769, 24691, 26279000 22279000 22279000 22278990, 22278996, 35696286, 22278996, 22278997, 22278998, 22278999, 264259, 29331824, 66714117, 29331825, 60432289, 35696052, 29331828, 264908, 66712502, 264512, 265007, 265008, 60170831,
	87775458 (4713, 4714) Novel Protein sim. GBank gil4929741[gb]AAD34131.1[AF15189 - (AF151894) CGI-136 protein [Homo sapiens]		UNCLASSIFIED	60432229, 60433356, 60433438, 264758, 52646317, 33109954, 21906764, 55811386, 87168474, 265017, 265018, 264605, 265019, 264681, 264682, 264768, 264686, 264768, 21906765, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906769, 265017, 260170615, 33657109, 27486264, 35695763, 55810764, 18108379, 35696423, 55811576, 35695855, 60170394, 56182323, 83373044, 18108385, 56526488, 264404, 60432113, 22279000, 264482, 264563, 264566, 264486, 264488, 264569, 18108394, 264259, 26331822, 18108370, 18108374, 264510, 265017, 264462, 264565, 264565, 264565, 264565, 264565, 264567, 265017, 265017, 26462, 264562, 264565, 264566, 264510, 265017, 265017, 26462, 264562, 264565, 264565, 264565, 264567, 265017, 265017, 265017, 265017, 26462, 264568, 264565, 264567, 26457,
	87777078 (4715, 4716) Novel Protein sim. GBank gil4218005 (AC006135) - putative vicilin storage protein (globulin-like) [Arabidopsis thaliana]		UNCLASSIFIED	264566, 264369, 18108354 22778997, 22278999, 264509, 264905, 264592, 18108351, 264681, 264682, 264769, 32833986, 18108374, 264556, 18108385,
I ~	87755859 (4717, 4718) Novel Protein sim. GBank gil1086830 (U41264) - coded for by C. elegans cDNA yk20f8.5, coded for by C. elegans cDNA yk491.5; coded for by C. elegans cDNA yk4891.5; coded for by C. elegans cDNA yk3696.5; coded for by C. elegans cDNA yk3696.5; coded for by C. elegans cDNA yk20f8.5; coded for by C. elegans cDNA yk20f8.5; coded for by C. yk16912		UNCLASSIFIED	264162 35696286, 22278998, 264905, 264511, 265007, 265008, 60433438, 264288, 264686, 21906769, 265020, 264692, 35695855, 264558, 56526486, 264563
	80046125 (4719, 4720) Novel Protein sim. GBank gij3881545 emb CAA93779 - (Z69904) cDNA EST yk428d5.3 comes from this gene; cDNA EST yk428d5.5 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	22278997, 29331826, 263981, 22279000
$\overline{}$	Novel Protein sim. GBank gil746487 (U23514) - No definition line found [Caenorhabditis elegans]			22278995, 22278999, 264512, 265009, 264757, 21906765, 65274620, 18108370, 60431528, 18108374, 264635, 60170394, 264482

						į
2362	91721193 (4723, 4724)	2362 91721193 (4723, 4724) Novel Protein sim. GBank		UNCLASSIFIED	22278999, 264259, 29331822, 29331824, 60432280, 264600, 264612, 60432229	
		CHAIN IB (MYOSIN HEAVY CHAIN IL)			60433356, 264448, 264682, 264683, 264369,	
					21906765, 21906768, 21906769, 60432113,	
3	_	100001			22279UUU, 22279UUZ	_
2363		95006635 (4725, 4726) Novel Protein sim. GBank gil854065 emb CAA58337 - (X83413) U88 Human herpesvirus 6]		UNCLASSIFIED	264907, 264629, 264635	
2364	94827104 (4727, 4728)	2364 94827104 (4727, 4728) Novel Protein sim. GBank	Contains protein domain (PF00389) - reductase	reductase	264488, 18108394, 264887, 18108398,	
		gi 5639830 gb AAD45886.1 AF14601 - (AF146018)	D-isomer specific 2-hydroxyacid		22278996, 56994075, 35696286, 22278997,	
		hydroxypyruvate reductase [Homo sapiens]	dehydrogenases		22278998, 264259, 66714117, 29331825,	_
					35696052, 264509, 264905, 264906, 264907,	
					264908, 66712502, 264909, 264511, 265006,	
					264512, 265007, 265008, 33657402, 264758,	_
					21906754, 87168474, 265010, 87168559,	
					264603, 265017, 265018, 265019, 264760,	
					264762, 18108351, 264448, 264764, 264683,	
					264684, 264288, 18108355, 264766,	
					18108358, 264689, 18108359, 21906765,	
					21906766, 21906767, 35695917, 265020,	
					265021, 265022, 60170615, 52644150,	
					264691, 33657023, 264692, 18108364,	_
					33657109, 18108368, 18108370, 18108374,	_
					35696423, 35695855, 264635, 264556,	
					264557, 264639, 60170394, 83373044,	
					18108383, 18108384, 18108385, 18108388,	
					56526486, 264482, 264564, 264486	-1
2365		94140746 (4729, 4730) Novel Protein sim. GBank gi 1840045 (U49082) -		transport	22278996, 22278998, 22278999, 264907,	
		transporter protein [Homo sapiens]			264909, 264910, 33657402, 264758, 264600,	
					264766, 264687, 264689, 21906765,	-
					21906767, 21906768, 21906769, 265021,	
					33657023, 33657109, 83373044, 264566	-
2366	94312388 (4731, 4732)			UNCLASSIFIED	52644507, 264259, 29331824, 29331825,	
					29331826, 29331827, 29331828, 264907,	
					29331830, 264909, 264511, 265008,	
					33657402, 264595, 52646317, 265017,	
					265018, 265019, 264605, 264685, 264766,	
_					264689, 21906766, 21906769, 35695917,	_
					265020, 265021, 265022, 52644150,	
					35695855, 52644332, 18108385, 18108387,	
					264564, 264566	
2367		94140910 (4733, 4734) Novel Protein sim. GBank gi 1065457 (U40410) - C54G7.4	Contains protein domain (PF00400) - UNCLASSIFIED	UNCLASSIFIED	35696286, 21906768, 55810764, 65274791,	
		gene product [Caenorhabditis elegans]	WD domain, G-beta repeat		264567	Т
2368	94322190 (4735, 4736)				264628	7

52644507, 52646842, 35696286, 264092, 264094, 52645080, 35696052, 264107, 29331830, 52644045, 265006, 265007, 265009, 52644296, 52644229, 264689, 21906765, 21906766, 35695917, 265020, 52644150, 263967, 36557109, 27486265, 35695763, 18108370, 263974, 18108376, 52664337, 263881, 18108376, 52664337, 263881, 18108376, 52664337, 263881, 18108378, 52664337, 263881, 18108378, 52664337, 263881, 18108378, 52664337, 263881, 18108378, 52664337, 263881, 18108378, 52664337, 263881, 18108388, 52664337, 263881, 18108388, 52664337, 263881, 18108388, 52664337, 263881, 18108388, 52664337, 263881, 18108388, 52664337, 263881, 18108388, 52664337, 263881, 18108388, 52664337, 263881, 18108388, 52664337, 263881, 18108388, 52664337, 263881, 18108388, 52664337, 5263881, 18108388, 52664337, 5268881, 526881, 5268881, 526881, 526881, 526881, 526881, 52688881, 5268881, 5268881, 5268881, 5268881, 5268881, 5268881, 526888	264508 264909 264596	264369	263067 263081	29331826, 265010, 265019, 35695917, 264634, 60432113	265006, 265007, 265008, 265009, 265011, 264766, 3695917, 35695855, 263981, 264557, 264557	60432049, 29331824, 264907, 52644045.	264512, 60433356, 21906754, 52644296, 87168559, 264448, 21906765, 21906768	21906769, 33657023, 18108368, 55811576,		60433438, 21906754, 265017, 265018,	265019, 264682, 264288, 264685, 21906767, 263972, 35695855, 87168518, 60432113	18108394 65274572 22278997 2227899	264095, 29331822, 29147620, 29331824,	66714117, 29331825, 29331826, 29331828,	33656970, 29146498, 29146499, 264509,	265006, 265007, 265008, 265009, 60170831,	265010, 265011, 265018, 55811150,	18108351, 264764, 264288, 21906767,	2265703 23657100 4040029 203021	18108379 35596423 264556 83373044	18108385, 18108388, 55526486, 22279000	22279002, 264563	65274572, 56182575, 35696286, 264259,	29331822, 29331824, 66714117, 29331825,	29331826, 60432289, 29331827, 56182435,	264510, 265009, 60433356, 87168474,	265011, 265018, 264288, 21906765,	33657023, 264557, 56182323, 83373044,	265017, 264288, 21906768	
struct	UNCLASSIFIED		UNCI ASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED			ATPase_associal			UNCLASSIFIED											UNCLASSIFIED						kinase	
					Contains protein domain (PF01138) - UNCLASSIFIED 3' exoribonuclease family				Contains protein domain (PF00628) - ATPase_associated	PHD-finger		Contains protein domain (PF01388) - UNCLASSIFIED	ARID DNA binding domain							,									Contains protein domain (PF00400) - kinase	אים טלייומווי, ט־טפומ ופאפמו
2369 94314334 (4737, 4738) Novel Protein sim. GBank gil5360901 dbj BAA82158.1 - (AB029343) a-helix colled-coll rod homologue [Homo saplens]				87418611 (4745, 4746) Novel Protein sim. GBank gil4589582 db BAA76813.1 - (AB023186) KIAA0969 protein [Homo sapiens]	Novel Protein sim. GBank gi 5105131 dbj BAA80445.1 - (AP000061) 246aa long hypothetical ribonuclease PH Aeropyrum pernix	87731355 (4749, 4750) Novel Protein sim. GBank	gi 1351115 sp P47758 SRPB_MOUSE - SIGNAL RECOGNITION PARTICLE RECEPTOR BETA SUBUNIT	(SR-BETA)	87613744 (4751, 4752) Novel Protein sim. GBank gi 2645435 (AF007780) - CHD3	[Drosophila melanogaster]		95319689 (4753, 4754) Novel Protein sim. GBank gi[5257005]gb AAD41239.1] -	(AF083249) Rb binding protein homolog [Homo sapiens]										9413/U32 (4/35, 4/3b)/Novel Protein sim. GBank gi[1072198 (U40942) - No	definition line found [Caenorhabditis elegans]					Novel Protein sim. GBank gi[3337357 (AC004481) - hypothetical protein [Arabidopsis thailana]	
94314334 (4737, 4738)	-7					87731355 (4749, 4750)						95319689 (4753, 4754)												-					65444324 (4757, 4758) Novel Protein sim. G	
2369	2370	2371	2372	2373	2374	2375			2376			2377										į	620						2379	

		86923062 (4759, 4760) Novel Protein sim. GBank gil4502939 ref NP_001845.1 pCOL1 - collagen, type XI, alpha 1	Contains protein domain (PF01410) - collagen Fibrillar collagen C-terminal domain	collagen	264908, 264910, 265011
2381		Novel Protein sim. GBank gij4455609jembjCAB36555j - (AL031846) dJ742C19.5 (novel Chromobox protein) [Homo sapiens]		helicase	56182575, 264091, 264093, 264259, 29331825, 264105, 264906, 60433356, 21906754, 265017, 265019, 264683, 264288, 264685, 264686, 264691, 264692, 264693, 55811576, 264636, 264567
2382		91225982 (4763, 4764) Novel Protein sim. GBank gil4325130lgb AAD17276j - (AF119716) dMi-2 protein [Drosophila melanogaster]	Contains protein domain (PF00628) - Iransport PHD-finger	transport	29331824, 60432289, 264905, 264596, 21906754, 264769, 265022, 264693, 263967, 33657109, 264629, 264631, 264558, 83373044, 60432113, 264482
2383	_	87442841 (4765, 4766) Novel Protein sim. GBank giļ1902982 dbj BAA19005 - [(D89049) lectin-like oxidized LDL receptor [Bos taurus]	Contains protein domain (PF00059) - glycoprotein Lectin C-type domain	glycoprotein	265009, 21906765, 21906766
2384		95354766 (4767, 4768) Novel Protein sim. GBank gil2462851 (AF016252) - Spinophilin [Rattus norvegicus]	Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF).	struct	264488, 52644507, 52645156, 52646365, 35696286, 22278999, 52645080, 29331824, 29331826, 35696052, 29331828, 264906, 264828, 52644045, 265006, 265008, 265009, 33109954, 33657084, 52644296, 265011,
					265017, 265018, 264683, 52644229, 21906785, 21906787, 21906786, 265020, 52644150, 33657023, 246893, 65274620, 52644150, 33657102, 246893, 652748261, 32685762, 27486264, 33657349, 27486265, 264555, 264556, 264555, 264556, 264555, 264555, 264556, 264555, 264556, 264557, 264556, 22579000, 22779002, 227790002, 22779002, 22779002, 22779002, 22779002, 22779002, 22779000
2385	95419485 (4769, 4770)			UNCLASSIFIED	264488, 52644507, 52645156, 264887, 52646365, 22278995, 22278996, 22278997, 22278999, 264259, 52645080, 29331822, 29331826, 35696052, 52644045, 265006, 265007, 265008, 265009, 264410, 6043229, 6043336, 5264631, 2496674, 265019, 21906765, 21906764, 266019, 26448, 26681, 264681, 264689, 25811957, 265021, 265022, 264690, 264691, 265021, 265022, 264690, 264631, 526454332, 22279000, 22279002, 264563, 264565, 264567
2386	94742649 (4771, 4772)	94742649 (4771, 4772) Novel Protein sim. GBank gil4929699[gb AAD34110.1 AF15187 - (AF151873) CGI-115 protein [Homo sapiens]		glycoprotein	264488, 22278995, 22278996, 22278997, 264259, 29146499, 264112, 264511, 60170831, 60432229, 264595, 60433438, 87168474, 87168559, 264682, 21906765, 21906766, 21906766, 21906767, 21906769, 29148629, 35695817, 265021, 264699, 33657109, 264628, 18108376, 284566, 284566, 284487

264634	264595	264488, 22278998, 22278999, 264509,	264905, 264906, 264907, 264908, 264909,	265006, 264511, 264512, 264910, 264591,	21906754, 264601, 264604, 264761,	18108351, 264764, 264288, 264766, 264768,	264769, 21906765, 21906768, 264692,	264693, 35696423, 264635, 264636, 264555,	83373044, 22279000, 264486	52644507, 56182575, 22278995, 35696286,	22278996, 22278997, 22278999, 29331822,	29331825, 29331826, 35696052, 264905,	52644045, 265009, 264758, 264759,	33109954, 52644296, 85658542, 265011,	265017, 265018, 264605, 52644229,	21906765, 21906767, 21906768, 21906769.	35695917, 52644150, 33657023, 33657109.	33657349, 35695763, 18108370, 18108374,	18108376, 35696423, 35695855, 264555,	52644332, 56182323, 60170394, 83373044,	56526486	263976	35696286, 35696052, 264508, 264905,	264509, 264906, 264907, 264908, 264909,	264510, 264511, 264512, 264910, 265009,	264591, 264758, 264600, 264604, 264762,	264448, 264764, 264369, 264766, 264768,	264769, 264689, 35695917, 264629,	18108374, 263978, 35696423, 35695855.	264631, 264634, 264635, 264636, 264637,	264639, 60170394, 264639, 264565, 264486
UNCLASSIFIED	UNCLASSIFIED	dna_rna_bind								polymerase												UNCLASSIFIED	nuclease								
		Contains protein domain (PF00170) - dna_ma_bind	bZiP transcription factor							Contains protein domain (PF00476) - polymerase	DNA polymerase family A												Contains protein domain (PF00560) - nuclease	Leucine Rich Repeat							
		Novel Protein sim. GBank	gil4758058 reflNP_004372.1 pCREB - cAMP responsive	element binding protein-like 1						Novel Protein sim. GBank gi[1644239 dbj BAA12223 -	(D84103) mitochondrial DNA polymerase gamma [Homo	sapiens											Novel Protein sim. GBank gi 4240169 dbj BAA74863.1 -	(AB020647) KIAA0840 protein [Homo sapiens]				-			
14997990 (4773, 4774)	11424604 (4775, 4776)	95310650 (4777, 4778) Novel Protein sim.								94320912 (4779, 4780) Novel Protein sim.												80036194 (4781, 4782)	94245016 (4783, 4784) Novel Protein sim.								
2387	2388	2389								2390						_		_				2391	2392		_						

3.1pRPLP - ribosomal protein, k gil4159888 (AC004908) - zinc of uncertain exon structure; similar 3) [Homo sapiens]
Contains protein (PTU0400) - Indosomal protein L10 Ribosomal protei
Contains protein (PTU04b0) - Inossomal protein L10 Ribosomal protei
Contains protein domain (PF00466) - irbosomalprot Ribosomal protein L10 Ribosomal protein
95302633 (4785, 4786) Novel Protein sim. CBank gil4159888 (AC004909) - zinc finger, CCH2 type to Q99276 (PID_93025333) [Homo saptens] 95237212 (4789, 4790) Novel Protein sim. GBank gil4159888 (AC004909) - zinc finger, CCH2 type to Q99676 (PID_93025333) [Homo saptens]
3.1pRPLP - ribosomal protein, fig. 1914159888 (AC004908) - zinc of uncertain exon structure; similar of uncertain exon structure; similar shows sapiens]
3.1pRPLP - ribosomal protein, fig. 1914159888 (AC004908) - zinc of uncertain exon structure; similar of uncertain exon structure; similar shows sapiens]
3.1pRPLP - ribosomal protein, fig. 1914159888 (AC004908) - zinc of uncertain exon structure; similar of uncertain exon structure; similar shows sapiens]
3.1pRPLP - ribosomal protein, fig. 1914159888 (AC004908) - zinc of uncertain exon structure; similar of uncertain exon structure; similar shows sapiens]
93 95302633 (4785, 4786) Novel Protein sim. GBank gil4506667[ref[NP_000993.1]pRPLP large, P0 large, P0 large, P0 large, P0 large, P0 linger protein sim. GBank gil41598 finger protein from gene of uncertal to Q99676 (PID:g3025333) [Homo: g15712726[gb]ACA7636.1]AF160
93 95302633 (4785, 4786) Nov gild ang g
93 95302633 (4785,
93 9530. 394 9432 395 9528
163 (2)

52646365, 18108397, 56182575, 35696286, 22278997, 22278999, 60432049, 264259, 29331824, 29331824, 29331824, 29331825, 29331826, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331837, 66712502, 264100, 60170831, 264591, 33657402, 60433438, 55812038, 33109954, 21906754, 33657402, 60433438, 56812038, 264288, 264766, 52644229, 21906766, 21906767, 265020, 265021, 60170615, 264692, 33657023, 6574460, 52645129, 33657182, 27486262, 27486264, 27486265, 264629, 18108374, 35696423, 35695855, 264631, 264565, 52644332, 264558, 2645482, 22279002, 264488	5264507, 52645156, 56182575, 264259, 29147620, 264905, 264907, 264908, 264909, 264910, 264758, 52644296, 264603, 264604, 264762, 264681, 264764, 18108357, 264769, 21906768, 264693, 264628, 264635, 264638, 264639, 264564		1906766, 52646842, 56994075, 33657182, 27486262, 52644296, 265017
Contains protein domain (PF00560) - nuclease Leucine Rich Repeat		UNCLASSIFIED	phosphatase
Novel Protein sim. GBank gi 106322 pir B34087 - hypothetical protein (L1H 3' region) - human		88047689 (4795, 4796) Novel Protein sim. GBank gil3258609 (AC005178) - H53_GS1 [Homo sapiens] B7738965 (4797, 4798) Novel Protein sim. GBank gil786117 (L41834) - nuclear protein [Ensis minor]	91214116 (4799, 4800) Novel Protein sim. GBank gil2352822 gb AAB69285.1 - (AF008945) glucose-6-phosphatase [Haplochromis nubilus]
396 95096700 (4791, 4792)			2400 91214116 (4799, 4800)

	,					_	
52644507, 52645156, 52644229, 264688, 21906764, 21906765, 52646365, 52646842, 21906766, 21906765, 52646365, 52646842, 21906766, 21906767, 21906768, 22278995, 356934075, 3569628, 22278996, 22278997, 266020, 22278998, 22278997, 266020, 22278998, 22278997, 266020, 22645080, 264693, 33657023, 52645080, 2646686, 33656970, 33657349, 27486265, 33656970, 33657349, 27486265, 35695652, 265007, 265008, 265009, 264637, 52644332, 55812038, 52646317, 18108385, 52644296, 87168474, 265010, 87168599, 60432113, 265017, 265018,	264907, 264908, 264909, 264566	22278999, 35696052, 265018, 264686, 264693, 83373044, 264567	35696286, 29331826, 35696052, 265008, 265018, 21906769, 264564	35696286, 29331822, 265007, 21906754, 265017, 265018, 265019, 264763, 264369, 21906765, 35695917, 265020, 265021, 52644150, 264693, 35695855, 264632, 52644332, 22279002	52646842, 22278994, 22278996, 35696286, 22278997, 22278999, 22278999, 264092, 264093, 60432049, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 264106, 264906, 55640465, 265007, 265008, 265009, 264910, 264592, 60433356, 60433438, 33109954, 265010, 265011, 265018, 265019, 264369, 264686, 21906768, 21906769, 2564150, 264632, 56182323, 264639, 22279002, 264539, 22279002, 264536, 2645369, 262279002, 264536, 2645369, 262279002, 264536, 264546,	264685, 264686	264910, 265010, 264448, 264557
phosphatase			dehydrogenase	UNCLASSIFIED	struct	UNCLASSIFIED	UNCLASSIFIED
		Contains protein domain (PF00062) - C-type lysozyme/alpha-lactalbumin family	Contains protein domain (PF00106) - dehydrogenase short chain dehydrogenase		Contains protein domain (PF01302) - struct		
91214118 (4801, 4802) Novel Protein sim. GBank gil2352822lgblAAB69285.1 - (AF008945) glucose-6-phosphatase [Haplochromis nubilus]	91221408 (4803, 4804) Novel Protein sim. GBank gil4689258 gb AAD27832.1 AF12185 - (AF121859) sorting nexin 9 [Homo sapiens]		milar	94311851 (4809, 4810) Novel Protein sim. GBank gi 464178 dbj BAA03581 - (D14853) polyprotein [Hepatitis C virus]	88094501 (4811, 4812) Novel Protein sim. GBank gi[2773363 (AF041362) - microtubule binding protein D-CLIP-190 [Drosophila melanogaster]		87391503 (4815, 4816) Novel Protein sim. GBank gil423442 pir S33513 - gene Fif protein - mouse
91214118 (4601, 4802)	91221408 (4803, 4804)	94135432 (4805, 4806)				_	
2401	2402	2403	2404	2405	2406	2407	2408

22278995, 22278996, 22278997, 264097, 29331822, 29331824, 29331827, 29146498, 52644045, 60433438, 33657084, 87168474, 264760, 21906767, 29148627, 29148629, 52644150, 33657023, 263967, 20281069, 18108374, 20281071, 56182323, 83373044, 18108385, 87168518	22278998, 264259	29331824, 29331827, 29331828, 264764, 264369, 33657109, 56182323	22278998, 29331827, 264907, 265011, 265017, 265018, 265019, 18108351, 21906766, 265020, 33657109, 264559, 18108385	264488, 264259, 29331826, 264508, 264905, 264509, 264907, 264908, 264510, 264511, 264512, 265008, 265009, 264910, 265011, 264682, 264764, 264766, 264686, 264768, 264768, 264689, 264768, 39657023, 18108370, 264628, 36595855, 264632, 264634, 264635, 264566, 264567, 26457, 264567, 264567, 264567, 264567, 264567, 264567, 264567, 264567, 264567, 26457, 26	52645156. 52646342. 52646385. 18108398. 56182575, 22278994. 22278995, 56994075, 22278996. 35696286. 22278997, 22278998. 22278999. 264259. 52645080. 29331822. 22331824. 29331825. 29331828. 26331824. 29331825. 29331826. 29331827. 26331827. 26331828. 33657402. 264594. 264603. 6433356. 33657402. 265019. 264369. 2646317. 21906764. 3365708. 265019. 264369. 21906764. 264688. 56181562. 21906767. 29148627. 21906765. 21906766. 2190677. 29148627. 21906769. 264693. 265019. 264638. 264638. 264639. 265019. 264638. 2646182. 27486262. 33657109. 33657109. 33657109. 33657109. 264638. 2
UNCLASSIFIED		ļuj	dna_rna_bind	struct	ubiquitin
	Contains protein domain (PF00080) - Copper/zinc superoxide dismutase (SODC)		Contains protein domain (PF00076) - dina_rna_bind RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	Contains protein domain (PF00038) - struct Intermediate filament proteins	
2409 94741770 (4817, 4818) Novel Protein sim. GBank gi 1176601 sp P45966 YNZ6_CAEEL - HYPOTHETICAL 20.8 KD PROTEIN T09A5.6 IN CHROMOSOME III	87604860 (4819, 4820) Novel Protein sim. GBank gil4966262lgb AAC48052.2 - (U64849) Contains similarity to Pfam domain: PF00646 (F-box), Score=28.7, E-value=4.3e-05, N=1 [Caenorhabdilis elegans]	87534633 (4821, 4822) Novel Protein sim. GBank gij3114713 (AF061346) - Edp1 protein [Mus musculus]	87778332 (4823, 4824) Novel Protein sim. GBank gij5410336jgb AAD43038.1 - (AF106685) myelin gene expression factor 2 [Homo sapiens]	94133820 (4825, 4826) Novel Protein sim. GBank gij5262705 emb CAB45778.1	94312590 (4827, 4828) Novei Protein sim. GBank gij1082340 pir S52863 - DNA-binding protein R kappa B - human
7, 4818) N	4820) A	1, 4822) N	3, 4824) N	(1) (1) (1) (1) (1) (1) (1) (1) (1) (1)	7, 4828) N
94741770 (481	2410 87604860 (4819,	87534633 (482	2412 87778332 (482	2413 94133820 (48	2414 94312590 (482

264259, 264908, 60433356, 33657402, 21906754, 265018, 264687, 264689, 21906769, 55811957, 265021, 264690, 264691, 33657023, 264693, 35696423, 56182323, 56526486	264638	264094, 29331822, 29331824, 29331827, 264369	56994075, 264091, 264259, 29331824, 29331825, 6432289, 29331822, 264907, 264917, 265009, 60432229, 21906754, 87168559, 265019, 264682, 21906768, 21906769, 265020, 265021, 23557023, 6277420, 18108385, 22279000, 22279002, 264566	264569, 264762, 264448, 264691, 264631, 264634, 264555, 264556, 264638, 264558	29331824, 29331825, 29331828, 60432229, 33109954, 85658542, 87168474, 265018, 264288, 265020, 264564	264909, 264768, 264638	18108392, 18108394, 18108398, 264906, 265006, 265010, 18108351, 18108374, 18108385	264259, 60432049, 29331822, 29331826, 60432289, 29331822, 26508, 265009, 6043336, 21906754, 265017, 265018, 265019, 21906766, 21906768, 21906769, 265020, 265021, 20281149, 263971, 6043213	56182575, 35696286, 22278997, 60432049, 264259, 29331826, 29331828, 264905, 66712502, 29331830, 60433356, 265011, 265019, 264766, 21906768, 55811957, 264692, 33657023, 33657109, 55811576, 56182323, 83373044, 18108385, 18108388, 60432113, 22279000	264634	22278996, 22278998, 264259, 264102, 264512, 265006, 21906767, 18108370, 18108374, 263976
struct			UNCLASSIFIED	struct	ÜNCLASSIFIED	UNCLASSIFIED	transport	kinase	UNCLASSIFIED	kinase	UNCLASSIFIED
Contains protein domain (PF00063) - Myosin head (motor domain)		Contains protein domain (PF00153) - Mitochondrial carrier proteins	Contains protein domain (PF00411) - UNCLASSIFIED Ribosomal protein S11				Contains protein domain (PF00617) - transport RasGEF domain	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	·	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	
2415 88089002 (4829, 4830) Novel Protein sim. GBank gil423915 pir A45439 - myosin Contains protein domain (PF00063) - struct Myosin head (motor domain)	94118356 (4831, 4832) Novel Protein sim. GBank gij3025445 (AC004528) - R32184 1 [Homo sapiens]	87733334 (4833, 4834) Novel Protein sim. GBank gij 1084944 ptr 554495 - hypothetical protein YPR021c - yeast (Saccharomyces cerevisiae)	94234349 (4835, 4836) Novel Protein sim. GBank gi[1176572 sp P45895 YNA4_CAEEL - HYPOTHETICAL 91.0 KD PROTEIN PAR2.4 IN CHROMOSOME III	82374249 (4837, 4838) Novel Protein sim. GBank gij284006[pir S18732 - autoantigen, 64K - human	94844244 (4839, 4840) Novel Protein sim. GBank gi 1076211 pir 550755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii		88084714 (4843, 4844) Novel Protein sim. GBank gi[2224567 dbj BAA20772 - (AB002311) KIAA0313 [Homo sapiens]	Novel Protein sim. GBank gi 4505153 ref NP_002392.1 pMEKK - MAP/ERK kinase kinase 3	94854047 (4847, 4848) Novel Protein sim. GBank gi 2988398 (AC004381) - Unknown gene product [Homo sapiens]	87415981 (4849, 4850) Novet Protein sim. GBank gij2077932 dbj BAA19879 - IO86556) Protein Kinase (Rattus norveoicus)	87613945 (4851, 4852) Novel Protein sim. GBank gij2039388igb AAB53003.1 - (U94619) circulating cathodic antigen [Schistosoma mansoni]
88089002 (4829, 4830)	94118356 (4831, 4832)	87733334 (4833, 4834)	94234349 (4835, 4836)	82374249 (4837, 4838)	94844244 (4839, 4840)	87805345 (4841, 4842)	88084714 (4843, 4844)	2423 88058390 (4845, 4846) Novel Protein sim. G gil4505153 ref NP_0 kinase 3		_	
2415	2416	2417	2418	2419	2420	2421	2422	2423	2424	2425	2426

264594, 265011, 264760, 18108351, 264682, 265011, 264760, 18108351, 264682, 264683, 264369, 264684, 264686, 264687, 264689, 21906766, 264691, 264693, 18108374, 18108377, 264557, 264639, 18108385	22278996, 22278999, 35696052, 21906754, 264288, 21906765, 21906768, 21906769, 35695917, 265020, 263972, 22279002	22278999, 29331824, 264906, 264909, 264511, 265009, 21906754, 265017, 265018, 265019, 264448, 264683, 264288, 21906765, 21906768, 265021, 264693, 18108381	264112, 264691	29331826, 29331827, 35696052, 29146499, 264905, 264906, 264681, 264288, 264689, 21906765, 264692, 35696423	264634, 264558	264569, 264905, 265018, 264762, 264683, 264691, 264556, 264557, 264639, 264558	264563	264555	265008, 264758, 265010, 264689, 27486261, 263972, 18108374, 18108381	29331828, 265007, 265009, 265017, 264760, 264685, 264693, 264565	29331826, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264511, 265006, 264512, 264910, 265009, 264511, 265006, 264512, 264910, 265009, 264501, 33657402, 264686, 264761, 264761, 33657023, 264693, 264512, 264563, 264564, 264565, 264566, 264567
ribosomalprot		UNCLASSIFIED	UNCLASSIFIED		protease	transcriptfactor		UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED
Contains protein domain (PF00573) - ribosomalprot Ribosomal protein L4/L1 family		Contains protein domain (PF00711) - UNCLASSIFIED Beta defensins		Contains protein domain (PF01728) - FISJ cell division protein		Contains protein domain (PF00096) - transcriptfactor Zinc finger, C2H2 type					Contains protein domain (PF01256) - UNCLASSIFIED Uncharacterized protein family UPF0031
2) CGI-28	85732889 (4855, 4856) Novel Protein sim. GBank gi 1537070 (U63840) - nucleoporin p54 [Rattus norvegicus]	87769276 (4857, 4858) Novel Protein sim. CBank gi[601931 (M94316) - neurofilament-H [Oryctolagus cuniculus]	(098)		80083033 (4863, 4864) Novel Protein sim. GBank gij3876367 emb[CAA93287] - (269360) Weak similarity to Eimeria thrombospondin (PIR Acc. No. A45517); cDNA EST EMBL: M89266 comes from this gene; cDNA EST yk295b9.5 comes from Ithis gene [Caenorhabditis elegans]	k gi 2224593 dbj BAA20784 - 1omo sapiens		(970)	88044008 (4871, 4872) Novel Protein sim. GBank gi 1263289 (U47856) - fibroin-4 [Araneus diadematus]	putative	94143473 (4875, 4876) Novel Protein sim. GBank gij3860014 (AF091088) - unknown [Homo sapiens]
27 87622693 (4853, 4	2428 85732889 (4855, 46	2429 87769276 (4857, 48	30 86948827 (4859, 4860)	2431 87649884 (4861, 48	2432 80083033 (4863, 4)	2433 80055092 (4865, 4	2434 19520148 (4867, 4868)	2435 20759044 (4869, 4870)		2437 83363424 (4873, 4)	2438 94143473 (4875, 4)

2439	2439 94850650 (4877, 4878) Novel Protein sim. (AC004044) small r thaliana]	Novel Protein sim. GBank gil4263519[gblAAD15345] - (AC004044) small nuclear riboprotein Sm-D1 [Arabidopsis thaliana]	Contains protein domain (PF01423) - UNCLASSIFIED Sm protein	UNCLASSIFIED	60424179, 18108397, 56182575, 56181686, 56994075, 22278996, 35696286, 22278997, 22278999, 264259, 52645080, 29331822, 26182181, 29331822, 56182181, 29331822, 60432289, 29331826, 26331822, 29331828, 369805, 2946498, 264509, 264905, 264906, 2544045, 60431735, 264905, 264906, 2544045, 60431735, 265019, 18108351, 264448, 264288, 264768, 5264429, 56181562, 21906764, 21906765, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21486262, 27486262, 27486262, 27486262, 33657109, 33659763, 18108370, 60431850, 56810764, 35696423, 65274791, 32279000, 22279002, 264567
2440	87641733 (4879, 488			UNCLASSIFIED	29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264907, 264510, 265018, 265019, 265019, 265019, 265033, 264539, 22279002
2441	87623914 (4881, 488	2) Novel Protein sim. GBank gil3024889 sp P56524 Y288_HUMAN - HYPOTHETICAL PROTEIN KIAA0288 (HA6116)	7	UNCLASSIFIED	264488, 264629, 18108374, 264564
2442	87273590 (4883, 4884) Novel Protein sim. 914506013 ref NP_ 17, regulatory subun	Novel Protein sim. GBank gil4506013 ref NP_002703.1 pPPP1 - protein phosphatase 1, regulatory subunit 7	Contains protein domain (PF00560) - UNCLASSIFIED Leucine Rich Repeat	UNCLASSIFIED	22278996, 2227899, 29331822, 264768, 264693
2443	84305949 (4885, 4886)	2443 84305949 (4885, 4886) Novel Protein sim. GBank gil1170658 sp Q02975 KID1_RAT Contains protein domain (PF01352) - Itanscriptfactor - RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17)	Contains protein domain (PF01352) - It KRAB box	transcriptfactor	264906
2444	Ī	Novel Protein sim, GBank gil4758824lrefiNP_004280.1 pNRF3 - nuclear factor (erythroid-derived 2)-like 3	Contains protein domain (PF00170) - transcriptfactor bZIP transcription factor	transcriptfactor	264259, 18108382, 18108383, 18108385, 22279000
2445					264259, 35696052, 264369, 18108361
2446	-	Novel Protein sim. GBank gi 4753887 emb CAA05409.2 - (AJ002424) p65 protein [Rattus norvegicus]	Contains protein domain (PF00095) - proteaseinhib WAP-type (Whey Acidic Protein) four-disulfide core*	proteaseinhib	265011, 264689, 33657023, 263981, 18108385
2447	94845149 (4893, 4894) Novel Protein sim. gil4885613 ref NP_ tumorigenicity 5	Novel Protein sim. GBank gil4885613[ref]NP_005409.1[pST5] - suppression of tumorigenicity 5		cadherin	56182575, 264259, 29331824, 29331825, 29331827, 60433356, 60433438, 264758, 265018, 264692, 65274620, 60431528, 65274791, 56182323

2448	2448 R7749680 (4895 4896)			ſ	
: 				ONCLASSIFIED	ZZZ18996, ZZZ18991, ZZZ18999, Z9331826, 35696052, 264107, 264110, 87168474.
					87168559, 18108351, 21906767, 21906769
	_				27486262, 263976
2449		87869075 (4897, 4898) Novel Protein sim. GBank		cadherin	264259, 264828, 265007, 264595, 265021,
	_	gij728837(spjP39194/aLU7_HUMAN - !!!! ALU SUBFAMILY SQ WARNING ENTRY !!!!			56526486
2450				Ē	264906
2451		91014563 (4901, 4902) Novel Protein sim. GBank	Contains protein domain (PF00071) - divcoprotein		264093 29331822 29331824 29331825
		gi[1710021 sp P35290 RB24_MOUSE - RAS-RELATED	Ras family		66714117, 29331826, 29331828, 35696052,
		PROTEIN RAB-24 (RAB-16)			264907, 66712502, 29331830, 264910,
					265009, 264758, 265017, 265018, 264762,
					264448, 264288, 21906767, 265021,
					33657023, 264693, 33657109, 263969,
2452	1	91230509 (4903, 4904) Novel Protein sim GBank nil15040341dhilB4A132161		o de	83373044, 18108385 254469 254449 254599 252577 47408974
		(D86980) KIAA0227 [Homo sapiens]			204102, 204112, 204000, 2033/2, 101003/4, 83373044, 264563
2453			Contains protein domain (PF00560) - nofrecep		264509, 264512, 18108385
		to murine leucine-rich repeat protein; possible role in neural			
		development by protein-protein interactions; 93% similarity			
2454		Novel Protein sim Charles (1975)		1	
7		Sociosal (1907), 1909 Protein Sin. Chair Sin. Chair Sin (1908) - Extersin Contains protein domain (1700) - UNCLASSIFIED	Contains protein domain (PF00170)		263994, 66714117, 29331827, 264508,
		like protein - maize	bZIP transcription factor		264509, 264905, 264906, 264907, 264908,
					264909, 264510, 264511, 264512, 265009,
					264910, 264591, 264758, 264759, 265010,
					265011, 264603, 264604, 264760, 264761,
					264762, 18108351, 264764, 264765, 264766,
					264686, 264768, 264769, 264534, 264691,
_					264692, 33657023, 264693, 33657109,
					264628, 263978, 35695855, 264634, 264635,
					264637, 264638, 264639, 83373044,
3776	_				18108385, 264563, 264564, 264486
7433		832883U1 (48U3, 491U) Novel Protein Sim. GBank	Contains protein domain (PF00928) - glycoprotein		264488, 22278996, 264259, 35696052,
		giloscol (Splingsocial Andrew MOUSE - CLAIHKIN COAT	Adaptor complexes medium subunit		264905, 264906, 264907, 264908, 264909,
		ASSEMBLY PROTEIN AP47 (CLATHRIN COAT	family		264510, 264512, 265008, 265009, 264910,
		ASSOCIATED PROTEIN AP41) (GOLGI ADAPI OR AP-1			264591, 264592, 264593, 264594, 264758,
		4/ RD PROJEIN (HAT 4/ RD SUBUNIT) (CLATHRIN			265019, 264760, 264681, 18108351, 264683.
		ASSEMBLY PROTEIN ASSEMBLY PROTEIN COMPLEX 1			264764, 18108354, 264766, 264768, 264769,
		MEDIONI CHAIN)			264689, 21906766, 21906767, 21906769,
					29148629, 35695917, 265020, 265022,
				-	33657023, 33657109, 18108370, 264628,
					264629, 264631, 264632, 264635, 56182323,
					60170394, 18108385, 264563, 264564,
2456	88166700 (4011 4012)	Mana Despite the Contraction of			264566, 264567
<u> </u>	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Journal of Wall, valid, walk frotein smr. Ceank gill-Sebs30 (AC003079) - Ankyrin-Contains protein domain (PF00023) - Ikinase like; 54% similar to 20223404 (NID301233) in exons Spanning 40074 to 4454 of class o	-Contains protein domain (PF00023) - Ank repeat		264693
		spanning 43374 to 11331 of clone. [Homo sapiens]			

3/5 (4913, 4914)	2457 94118375 (4913, 4914) Novel Protein sim. GBank gi 3025447 (AC004528) - R32184_3 [Homo sapiens]		UNCLASSIFIED	56181686, 264905, 264907, 264511, 264596, 55811386, 264682, 264684, 264685, 264687,
				264691, 33657023, 264693, 35695855, 264636, 264555, 56182323, 264558, 56526486, 264563
15, 4916)	85875304 (4915, 4916) Novel Protein sim. GBank gil2384942 (AF022985) - Similar Ito collagen [Caenorhabditis elegans]		UNCLASSIFIED	264691, 264693, 264634, 264559
7, 4918)	1/AC00499 - (AC004997) AA538043 (NID:92284036)	Contains protein domain (PF00097) - Iransport Zinc finger, C3HC4 type (RING finger)	Iransport	27486265
19, 4920)	94315289 (4919, 4920) Novel Protein sim. GBank gil4929701[gb AAD34111.1 AF15187 - (AF151874) CGI-116 protein [Homo sapiens]		kinase	65274572, 35696286, 22278996, 22278997, 60432049, 29331822, 29331824, 29331825, 66714117, 60432289, 29331826, 29331827, 29331828, 33656970, 29146499, 264102, 264109, 60433438, 255017, 255018, 25619, 256288, 21906756, 21906766, 21906769, 35695917, 265020, 264691, 33657023, 27466261, 18108374, 35695855, 87168518, 60432113
21, 4922)	87645147 (4921, 4922) Novel Protein sim. GBank gil4426962 gblAAD20633 - (AF126062) Arf-like 2 binding protein BART1 [Homo sapiens]		UNCLASSIFIED	264259, 29331828, 264910, 18108351, 18108370, 18108374
23, 4924)	86998002 (4923, 4924) Novel Protein sim. GBank gij6420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]			264909, 264758, 264684, 18108374, 264637, 18108385
25, 4926)		Contains protein domain (PF00096) - UNCLASSIFIED Zinc finger, C2H2 type	UNCLASSIFIED	264681, 264566
27, 4928)	- I - I - I - I - I - I - I - I - I - I	Contains protein domain (PF01399) - protease	protease	264489, 52646842, 22278995, 35696286, 22278996, 22278999, 264259, 29331822, 29331824, 66714117, 29331825, 29331825, 29331827, 35696052, 29331826, 26509, 264509, 264501, 264511, 264509, 265008, 60170831, 264593, 52646317, 33109954, 33657084, 265017, 265018, 265019, 264762, 26448, 264764, 264288, 264766, 21906768, 21906769, 265021, 33657023, 33657109, 18108331, 60170394, 18108385, 22279002, 264488

	1	Ė –							
18108392, 52644507, 52645156, 52645365, 22278994, 22278995, 35696286, 22278996, 22278996, 22278996, 22278996, 22278998, 264259, 29331822, 29331824, 29331825, 60424269, 60432289, 29331827, 35696052, 29331828, 264907, 29331830, 25504045, 265009, 265009, 264910, 33657402, 60433438, 55812038, 21906754, 33109954, 265010, 265011, 18108351, 264369, 264288, 265018, 265019, 18108351, 264369, 264287, 21906765, 21906767, 21906768, 55811957, 35695917, 265020, 265021, 60170615, 52644150, 33657023, 35695763, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264568, 264638, 26433713, 264563, 264586, 264586, 264587, 26432113, 264563, 264638, 264371, 264563, 264638, 264371, 264563, 264587, 264587, 264588, 264387, 264563, 264587, 264568, 264487	264369	56994075, 264908, 21906768, 33657023	29331824, 52644045, 265008, 264910, 265019, 21906765, 21906769, 265021	264288, 264628	264592, 264692, 264555, 264556, 264557, 264558, 264559, 18108385, 264482	65274572, 35696286, 29331827, 265007, 264592, 33109954, 265018, 265019, 264288, 55811957, 265020, 264693, 55811576, 56182323	284092, 60432049, 29331825, 60433356, 265010, 265011, 18108351, 264764, 264288, 264692, 65274620, 18108370, 18108372, 18108374, 264634, 18108385	60424179, 65274572, 22278999, 60424269, 29331826, 265008, 60433356, 60433438, 265010, 18108351, 264448, 264288, 264687, 264689, 265021, 264692, 65274620, 60431528, 65274791, 264556, 56182323, 60432113	65274572, 56994075, 264259, 29331826, 60170831, 265017, 265018, 265019, 264683, 264369, 265020, 264693, 264564
олсодепе		UNCLASSIFIED	kinase		UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED		
Contains protein domain (PF00069) - oncogene			Contains protein domain (PF00400) - kinase WD domain, G-beta repeat						
Novel Protein sim. GBank gil4506401frefNP_002871.1 pRAF1 • v-raf-1 murine feukemla viral oncogene homolog 1	85681386 (4931, 4932) Novel Protein sim. GBank gil4321619JgblAAD15788.1] - (AF051098) seven transmembrane domain orphan receptor [[Mus musculus]	Novel Protein sim. GBank gi[3513300 (AC005595) - F16601_1, partial CDS [Homo sapiens]	87614696 (4935, 4936) Novel Protein sim. GBank gi 2143455 pir 158106 - gene DMR-N9 protein - mouse (fragment)		80223831 (4939, 4940) Novel Protein sim. GBank gi[5420389]emb[CAB46680.1] - [(AJ243460) proteophosphoglycan [Leishmania major]	91013681 (4941, 4942) Novel Protein sim. GBank gi[5419882 emb CAB46424.1 - (AL096749) DKFZp434G153 [Homo sapiens]	95060811 (4943, 4944) Novel Protein sim, GBank gil4929747[gb AAD34134,1[AF15189 - (AF151897) CGI-139 protein [Homo sapiens]	Novel Protein sim. GBank gi 4539009 emb CAB39630.1 - (AL049481) putative protein [Arabidopsis thaliana]	Novel Protein sim. GBank gi 3252827 (AC004382) - Unknown gene product [Homo sapiens]
2465 95357483 (4929, 4930) Novel Protein sim. GBank gil4506401frefINP_002877 feukemia viral oncogene h				-+					94315616 (4947, 4948) Novel Protein sim. G
246;	2466	2467	2468	2469	247(2471	2472	2473	2474

2475	94321693 (4949, 4950)	3Bank gi 1216486 (U48852) - HT protein	Contains protein domain (PF00008) -	tgf	264259, 29331822, 265006, 265007, 265010,
		[Cricetulus guseus]	EGF-like domain		265011, 264448, 264288, 264369, 264685,
					264686, 18108357, 264768, 18108362,
					264693, 18108370, 18108374, 18108379,
					35696423, 83373044, 18108383, 18108385,
_					264564, 264565, 264567
2476		94315618 (4951, 4952) Novel Protein sim. GBank gij3252827 (AC004382) -		UNCLASSIFIED	264259, 60424269, 66714117, 264905,
		Unknown gene product [Homo sapiens]			265006, 264511, 265008, 265009, 264758,
					265010, 265011, 18108351, 264681, 264369,
					264288, 264689, 21906767, 265020,
					18108374, 264639, 18108382, 83373044,
					18108385, 87168518
2477	20718974 (4953, 4954)			UNCLASSIFIED	263978
2478	17659165 (4955, 4956)			UNCLASSIFIED	265017
2479		94314569 (4957, 4958) Novel Protein sim. GBank gi 1644232 db BAA11082 -		tm7	56994075, 22278999, 21906754, 264682,
		(D67066) N-WASP [Bos taurus]			21906765
2480	95295605 (4959, 4960)			UNCLASSIFIED	264905, 264907, 264765
2481	94718481 (4961, 4962)	94718481 (4961, 4962) Novel Protein sim. GBank gil5689469 dbj BAA83018.1 -		collagen	65274572, 56182575, 22278997, 264094,
		(AB028989) KIAA1066 protein [Homo sapiens]			264259, 29331822, 29331824, 66714117,
-,-			-		29331827, 35696052, 264508, 264905,
					264906, 264907, 264908, 52644045, 264909,
					56182435, 265008, 264910, 33657402,
					55812038, 264758, 265010, 265011, 265017,
					265018, 264760, 264762, 18108351, 264764,
					264288, 264766, 264686, 264768, 21906768,
					55811957, 265020, 264691, 264692, 264693,
					264629, 55811576, 264630, 264634, 264635
					264636 264637 264556 264558 56182323
					83373044 60432113 22279002
2482		87393165 (4963 4964) Novel Protein sim GRank oil 321249 pir II S28407 - quanioe		LINCI ASSIEIED	23313341, 23431824 24341825 24331827
<u> </u>		micleotide-exchange activator CDC25 homolog - mouse			25231522; 25331527; 25331523; 25331521; 264508 264905 264509 264906 264907
					264908 264511 264591 264768 264693
				-	264631 264632 264636 264638 264639
					264563
2483	87731583 (4965, 4966)			UNCLASSIFIED	264488, 22278995, 264093, 264095,
					60432049, 60433356, 60433438, 264448,
					264288, 263967, 18108370, 18108385,
	_		_		18108388, 264482
2484	_	94187774 (4967, 4968) Novel Protein sim. GBank		kinase	264563
		gij728831 spjP39188 ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII			
2485		87786556 (4969, 4970) Novel Protein sim. GBank gil1185397 (U25281) - SH3		UNCLASSIFIED	22278995, 22278996, 22278997, 22278999,
		domain binding protein [Rattus norvegicus]			264259, 60432049, 29331824, 60432289,
					29331827, 265007, 264910, 264593, 264600,
_					264603, 264604, 265019, 264448, 264288,
					264685, 264686, 264769, 264689, 35695917,
					265022, 264692, 264693, 56182323

265017, 264555	22278996, 35696286, 22278997, 22278998, 22278999, 264092, 264259, 29331822, 35696052, 264106, 264269, 264907, 29331830, 264099, 265006, 264507, 29331830, 264099, 265006, 264511, 265009, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906769, 3569517, 255021, 265022, 264534, 33657023, 264692, 33657109, 263972, 18108378, 56526486, 22279000, 22279002, 264563, 264482, 264565,	264910, 264448, 264288, 264684, 264691, 264634	264686, 264693, 55811576, 22279002	264907, 265008, 22279002	264259, 29331826, 265008, 264762, 18108370, 18108376, 18108379	264488, 22278998, 22278999, 29331828, 264591, 33109954, 265017, 55811150, 21906764, 21906768, 264692, 60431528, 87166518, 60432113, 22279000
	glycoprotein	UNCLASSIFIED	complementrecept	UNCLASSIFIED	tm7	transcriptfactor
	Contains protein domain (PF00071) - Ras family		Contains protein domain (PF00084) - complementrecept Sushi domain (SCR repeat)			
Novel Protein sim. GBank gi[2662167]dbj[BAA23715] - (AB007903) KIAA0443 [Homo sapiens]	GBank gil464559 sp P35287 RB14_RAT		GBank gi[4886439 emb CAB43355.1 - hetical protein [Homo sapiens]	Novel Protein sim. GBank gi[2588624 (AC003083) - Rap2 Interacting protein-like; similar to U73941 (PID:g1916018) [Homo sapiens]	Novel Protein sim. GBank gi 728832 sp P39189 ALU2_HUMAN - !!!! ALU SUBFAMILY SB WARNING ENTRY !!!!	Novel Protein sim. GBank gij3355303 (AF001549) - Unknown gene product [Homo sapiens]
2486 87748978 (4971, 4972) Novel Protein sim. (AB007903) KIAAC	95343105 (4973, 4974) Novel Protein sim. RAS-RELATED PR	87652451 (4975, 4976)	82990585 (4977, 4978) Novel Protein sim. (AL050253) hypot	2490 88069609 (4979, 4980) Novel Protein sim. Interacting protein [Homo sapiens]	91242116 (4981, 4982)	95308202 (4983, 4984) Novel Protein sim. Unknown gene pro
2486	2487	2488	2489	2490	2491	2492

52644507, 52645156, 22278995, 56994075, 35996286, 22278998, 264259, 52645080, 29331824, 29331825, 66714117, 60432289, 29331824, 29331827, 35696052, 29331828, 264508, 264510, 264512, 33657402, 60433438, 21906754, 52644296, 87168474, 87168559, 264603, 264681, 264448, 264683, 264288, 264689, 5264229, 264689, 21906765, 21906765, 21906767, 21906767, 21906767, 21906767, 21906767, 21906767, 21906767, 21906767, 21906767, 21906767, 21906767, 21906767, 21906767, 265020, 265021, 2564150, 33657023, 264693, 33657182, 35695763, 35696423, 35695855, 52644332, 33657043, 18108387, 871688518, 2779407,	65274572, 56182575, 35696286, 22278996, 56994075, 22278997, 60432049, 264259, 29331822, 29331824, 29331826, 29331827, 35696052, 264905, 264906, 264907, 264908, 264910, 264891, 60433356, 264595, 264510, 264511, 265007, 264910, 264591, 60433356, 264595, 55812038, 264758, 264596, 87168474, 87168559, 264600, 264601, 264602, 265017, 264604, 265018, 264604, 265018, 264606, 265017, 264604, 265018, 264606, 265017, 264604, 265018, 264606, 265017, 264604, 265018, 264606, 265017, 264604, 265018, 264606, 265017, 264604, 265018, 264506, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906766, 21906767, 21906769, 264638, 264638, 264638, 264638, 264538, 264538, 264538, 264538, 80452813, 22279002, 364566	65274572, 264907, 56182435, 265007, 264592, 264760, 18108351, 264448, 264369, 264288, 264684, 264686, 55811957, 265021, 264692, 33657109, 263973, 55811576, 264635, 264556, 264557, 264557, 264559, 37168518, 264563, 264487	264905, 264907, 264512, 265008, 265011, 18108351, 264448, 264288, 29148627, 264693, 18108370, 18108374, 18108385
Transport	dna_rna_bind	UNCLASSIFIED	
Contains protein domain (PF00459) - transport Inositol monophosphatase family	Contains protein domain (PF00651) - dna_rna_bind BTB/POZ domain	Contains protein domain (PF01237) - UNCLASSIFIED Oxysterol-binding protein	
Novel Protein sim. GBank gij3881275 emb CAA21725 - (AL032655) predicted using Genefinder; similar to Inositol monophosphafase family; cDNA EST yk255e11.5 comes from this gene [Caenorhabditis elegans]	Novel Protein sim. GBank gil4929615[gb AAD34068.1 AF15183 - (AF151831) CGI-73 protein [Homo sapiens]	Bank gil3041847 (AC004542) - NG PROTEIN-like; similar to P22059 o sapiens]	Novel Protein sim. GBank gi[2196874 emb CAA72538] - (Y11896) BRX protein [Mus musculus]
2500 94649324 (4999, 5000) Novel Protein sim. (AL032655) predict (AL032655) predict monophosphatase from this gene [Cae	94303896 (5001, 5002) Novel Protein sim. (gil4929615[gb AAD) protein [Hamo sapid	90993716 (5003, 5004) N	87878345 (5005, 5006) Novel Protein sim. G (Y11896) BRX protei
2500	2501	2502	2503

65274572, 22278994, 22278996, 22278998, 22278998, 60432049, 264259, 29331822, 29331824, 29331825, 66714117, 60432289, 29331826, 33656970, 264908, 66712502, 265007, 264910, 60170831, 6043229, 60433356, 60433438, 21906754, 87168474, 265017, 265018, 26448, 264288, 21906767, 256019, 265022, 264691, 33657023, 264693, 33657482, 27486262, 33657482, 27486262, 264827, 3657349, 18108370, 3569585, 264555, 5618232, 33657349, 18108370, 3569585, 264555, 2618223, 83373044, 60432113, 22279002	265017, 21906764, 265020, 264692	22278999, 264259, 29331825, 29331826, 29146499, 264907, 264909, 265006, 265008, 264591, 60432229, 21906754, 264763, 264683, 264766, 18108357, 264689, 21906769, 264693, 18108370, 263972, 18108374, 264558, 22279000	60424179, 52645156, 18108394, 22278994, 35696286, 66994075, 22278996, 29331822, 29331824, 60424269, 29331825, 29331827, 3365970, 60431735, 33657084, 87168559, 265017, 264448, 264369, 56181562, 21906766, 21906767, 21906768, 21906766, 21906767, 21906762, 21906766, 21906767, 21906762, 21906763, 1810836, 33657109, 274862261, 27486262, 33657349, 18108374, 55810764, 35696423, 56182323, 264558, 18108386	264510		264259, 29146498, 264905, 264288, 29148629, 35695917, 27486261, 264634	264091, 29331824, 29331825, 29331826, 29331828, 35696052, 264508, 264509, 264905, 264907, 264908, 264909, 264910, 33657402, 264757, 33109564, 265017, 265018, 264605, 264760, 264762, 264763, 264766, 264628, 55811576, 35696423, 264634, 264637, 264638, 264639, 87168518, 22279002, 264564
dna_ma_bind		transport	UNCLASSIFIED	UNCLASSIFIED	tm7	UNCLASSIFIED	UNCLASSIFIED
		Contains protein domain (PF00097) - transport Zinc finger, C3HC4 type (RING finger)			Contains protein domain (PF00001) - tm7 7 transmembrane receptor (rhodopsin family)		
2510 95421379 (5019, 5020) Novel Protein sim. GBank gij3293537 gb AAC25762.1 - (AF071059) zinc finger RNA binding protein [Mus musculus]	Novel Protein sim. GBank gil4323152[gb AAD16228.1] - (AF098863) Ets-protein Spi-C [Mus musculus]	88084771 (5023, 5024) Novel Protein sim. GBank gil4502075 ref NP_001135.1 pAMFR - autocrine motility factor receptor	95357843 (5025, 5026) Novel Protein sim. GBank gij3004657 (AF017777) - bobby sox [Drosophila melanogaster]	88094578 (5027, 5028) Novel Protein sim. GBank gi 2258437 (AF008197) - syncollin [Rattus norvegicus]	87994509 (5029, 5030) Novel Protein sim. GBank gij3757727lembjCAA18783j - (AL022727) dJ80119.7 (olfactory receptor-like protein ((hs6M1-3)) [Homo sapiens]		Novel Protein sim. GBank gil4220527lemb CAA23000 - (AL035356) putative protein [Arabidopsis thaliana]
0 95421379 (5019, 5020)	87384281 (5021, 5022)					6 87786908 (5031, 5032)	87784966 (5033, 5034)
251	2511	2512	2513	2514	2515	2516	2517

35696286, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 35696052, 29331828, 264907, 264909, 264511, 265007, 60432229, 60433356, 60433438, 55812038, 265010, 265017, 264448, 264288, 264689, 21906769, 265022, 52644150, 264693, 18108370, 263972, 264555, 56183233, 833373044, 18108385, 60432113, 224088	264259, 66714117, 29331826, 29331827, 29331828, 264907, 66712502, 265006, 265008, 265008, 265010, 265011, 265018, 264288, 21906769, 265020, 60431528, 55811576, 65274791, 264632, 264555, 264636, 22279002, 264564	264259, 264908, 264910, 264682, 21906769, 265020, 264563	264488, 264489, 263994, 65274572, 22278895, 22278898, 264259, 29331822, 22378895, 264269, 29331822, 28331822, 284307, 66712502, 264906, 264907, 66712502, 264591, 265007, 264591, 264592, 264594, 264594, 264595, 264595, 264682, 264681, 264488, 264763, 264682, 264686, 21906769, 55811957, 264692, 264693, 27486261, 18108370, 264692, 264629, 18108374, 55811576, 35696423, 35695855, 264566, 264563, 264563, 264563, 264564, 264565, 264565, 264565, 264567, 264563, 264564, 264565, 264567, 264563, 264567, 264564, 264565, 264567, 2645	264489, 22278997, 20281171, 21906754, 35695917, 263967, 263976, 263981, 20281169	263969 56994075, 22278996, 35696286, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331826, 29331828, 29331830, 56182435, 264512, 265008, 60170831, 33657402, 265001, 87168559, 265019, 264288, 21906765, 21906769, 35695917, 265020, 265021, 265022, 52644150, 264691, 33657023, 33657109, 27486261, 35696423, 65274791, 264559, 83373044, 56526486,
UNCLASSIFIED	kinase	transport	histone	transport	
Contains protein domain (PFD0018) - UNCLASSIFIED SH3 domain			Contains protein domain (PF00850) - histone Histone deacetylase family		Contains protein domain (PF01753) -
1 AF15181 - (AF151819) CGI-61	Novel Protein sim. GBank gi 4263748 gb AAD15420 - (AC004883) similar to KIAA0766; similar to PID:g3882253 [Homo sapiens]	Novel Protein sim. GBank gil4826722 ref NP_005085.1 pFATP - fatty acid transport protein 4	Novel Protein sim. GBank gij5174488jref NP_006035.1 pKIAA - histone deacetylase 6 Histone deacetylase family	Novel Protein sim. GBank gil45800111gb AAD24201.1 U81002 (U81002) TRAF4 associated factor 1 [Homo sapiens]	Novel Protein sim. GBank gi 1809327 (U76374) - skm- BOP2 [Mus musculus]
2518 94147410 (5035, 5036) Novel Protein sim. GBank gil4929591[gb AAD34056. protein [Homo sapiens]	2519 94326180 (5037, 5038) Novel Protein sim. (AC004883) simila [Homo sapiens]	2520 87413235 (5039, 5040) Novel Protein sim. gil4826722 ref NP_ protein 4	2521 95316244 (5041, 5042) Novel Protein sim. gi 5174489 ref NP_	87754052 (5043, 5044)	2523 95340467 (5045, 5046) 2524 95340469 (5047, 5048) Novel Protein sim. BOP2 (Mus muscu

2525 94126928 (5049, 5050) Novel Protein sim. GBank gi 2073564 (U80223) - eukaryotic
Novel Protein sim. GBank gil4589628 dbj BAA76836.1 - (AB023209) KIAA0992 protein [Homo sapiens]
88094580 (5053, 5054) Novel Protein sim. GBank gi[2258437 (AF008197) - syncollin [Rattus norvegicus]
88078380 (5055, 5056) Novel Protein sim. GBank gij2085786 (AC002086) - similar to zinc finger 5 protein from Gallus gallus, U51640 (PID:g1399185) [Homo sapiens]
86670926 (5057, 5058) Novel Protein sim. GBank gi[3786433 (AF098505) - similar to Arabidopsis thaliana male sterility protein 2 (SW:Q08891) [Caenorhabditis elegans]
87419778 (5063, 5064) Novel Protein sim. GBank gi[2664625 emb CAA16972 [(AL021811) putative protein [Arabidopsis thaliana]
87000255 (5065, 5066) Novel Protein sim. GBank gil437181 (U02289) - GTPase- activating protein [Caenorhabditis elegans]

2534	87332322 (5067, 5068)	2534 [87332322 (5067, 5068) Novel Protein sim GBank oii3452473 (AF084204)	CHICANICAL	204060 2000000 0000000000000000000000000
		Serine/threonine protein kinase TAO1 (Rattus norveoleus)		2406250 255020 255020 20017,
				21900709, 200020, 200022, 33057109,
2535	91225056 (5069, 5070) Novel Protein sim. Gl	Novel Protein sim. GBank gil4468311lemblCAB37992i -		CENTACTA DEPOSOR PRANCOS CONTROL
		(AL031432) dJ465N24.2.1 (PUTATIVE novel protein)		62712603 25506 50422230 25504
		lisoform 1) (Homo sanians)		00/120/2, 2000/06, 6U43ZZZB, 20001/,
				265018, 265019, 264288, 264369, 264689,
				21906768, 265020, 265021, 264636,
2	2000			60170394, 22279002
0507	2330 9421834U (3071, 3072) Novel Protein sim. G	Novel Protein sim. GBank	kinase	18108398, 56182575, 35696286, 22278997,
		gi/728836 sp P39193 ALU6_HUMAN - !!!! ALU SUBFAMILY		22278999, 60432049, 264259, 29331824.
		SP WARNING ENTRY IIII		29331826, 29331827, 29331828, 264905,
			_	264511, 265009, 264910, 264596, 52646317,
				18108351, 264681, 264683, 18108354,
				264288, 264687, 264769, 264689, 21905765,
				21906766, 21906767, 265021, 52645129,
				33657109, 18108374, 18108380, 56182323,
				18108381, 18108388, 87168518, 60432113,
2030	יי דיסי סביסיי מסמממיי			22279000, 22279002, 264567, 18108391
7663	93422203 (3073, 3074) Novel Protein sim. Gi	Novel Protein sim. GBank	ubiquitin	65274572, 35696286, 29331822, 29331825,
		gil455/026fretINP_003913.1[pHERC - guanine nucleotide		29331827, 29331828, 35696052, 264906,
		exchange factor p532		66712502, 264909, 265008, 265011, 264760
				264288 264685 35695917 60170615
				264691 33647023 64274620 33647100
				1010001, 00001,000, 0001,100, 00001,100,
			_	18108374, 35696423, 35695855, 264636,
2538	36853464 (6076 5076)			264558, 60170394, 56182323, 83373044
2620	-		UNCLASSIFIED	
6003	(0/00 ///00) 0/644146		UNCLASSIFIED	22278996, 22278999, 29331822, 29331825,
				29331828, 29146499, 264908, 264112,
				60170831, 87168559, 264604, 265019,
_				264685, 264766, 87168518, 22279000,
3	10000			264565, 264566
240	94218545 (5079, 5080)	94216343 (3079, 5080) Novel Protein sim. GBank gij1362647 pirl S53876 - sex-	UNCLASSIFIED	22278997, 29331828, 265008, 265009,
		regulated protein janus A - fruit fly (Drosophila		264758, 265010, 18108351, 264683, 264288,
		pseudoobscura)		21906765, 35695917, 265020, 18108374,
3	200000000000000000000000000000000000000			264567
1407	95308238 (5081, 5082)	95308238 (5081, 5082) Novel Protein sim. GBank	UNCLASSIFIED	264488, 264768, 264689, 264511, 20281171,
_		g 171103d sp 734/94/110_MOUSE - SER/THR-RICH		264634, 264635, 264691, 264639, 29331824,
		PROTEIN 110 IN DGCR REGION		264603, 264604, 264905, 264907, 264908.
				264766

2542	95298162 (5083, 5084)	2542 95298162 (5083, 5084) Novel Protein sim. GBank Ini15225320InblaAD40850 11AE08310 - (AE083107) simin	Contains protein domain (PF00220) - UNCLASSIFIED Naturabuscial hormones N	UNCLASSIFIED	264488, 18108394, 52646365, 52646842,	
		type 2 [Homo sapiens]	terminal Domain		002/10/2/2/2/10307, 00000200, 222/0030, (264259 52645080 29331822 29331824	_
					29331827, 35696052, 33656970, 264907,	
					264909, 52644045, 264510, 265006, 264512,	
					265007, 265008, 265009, 264910, 60431735,	
					52646317, 52644296, 265010, 265011,	
					265018, 265019, 18108351, 264683, 264288,	
					264685, 264687, 52644229, 264769,	
					21906766, 21906767, 21906769, 52644150,	
					33657023, 33657109, 52645129, 33657182,	
					27486261, 27486264, 33657349, 35695763,	
					18108374, 35696423, 35695855, 264631,	
-					264634, 264635, 264558, 83373044,	
_					18108385, 18108387, 87168518, 264563,	_
					264564	
2543	94139088 (5085, 5086)	94139088 (5085, 5086) Novel Protein sim. GBank gij5419857 jemb[CAB46374.1] -	Contains protein domain (PF00076) -		65274572, 56182575, 22278999, 264259,	
		(AL096723) hypothetical protein [Homo sapiens]	RNA recognition motif. (a.k.a. RRM.		29331826, 264907, 264510, 264511, 264592,	
			RBD, or RNP domain)		264595, 264764, 264369, 264288, 264684,	_
					264766, 264689, 21906765, 21906767.	
					21906769, 60170615, 264692, 264693,	
					55811576, 65274791, 264636, 264556,	
					18108381, 60170394, 264639, 18108385,	
					60432113, 22279000	
2544	94218549 (5087, 5088)	94218549 (5087, 5088) Novel Protein sim. GBank	Contains protein domain (PF00629) - glycoprotein	glycoprotein	18108397, 52646365, 22278997, 264259,	
		gil2498110 sp Q63191 AEGP_RAT - APICAL ENDOSOMAL MAM domain	MAM domain.		60432049, 29331822, 29331825, 29331826.	
		GLYCOPROTEIN PRECURSOR			29331827, 29331828, 264905, 264908,	
					265006, 265007, 265008, 87168559, 265017,	
_					265018, 265019, 18108351, 264448, 264686,	
					264687, 264689, 21906765, 265020, 265021,	
					18108370, 18108374, 18108376, 18108381,	
					18108385, 18108387, 56526486, 22279000,	
					264482, 264563, 264567	
2545	87742645 (5089, 5090)	2545 87742645 (5089, 5090) Novel Protein sim. GBank gij3327046 dbj BAA31591 -			29331825, 264906, 265009, 60170831,	
		(AB014516) KIAA0616 protein [Homo sapiens]			265017, 264369, 21906767, 60170615,	
25.30	000000000000000000000000000000000000000	4			264692, 33657109	寸
9407	86093601 (5091, 5092)	66093661 (5091, 5092) Novel Protein sim. GBank gil2996032 (AF054586) - brain	Contains protein domain (PF00097) - UNCLASSIFIED	UNCLASSIFIED	29331824, 265007, 22279002	
		inger protein [Rattus norvegicus]	Zinc finger, C3HC4 type (RING			
			liilgei			_

		gil4929607[gb]AAD34064.1[AF15182 - (AF151827) CGI-69 protein [Homo sapiens]	Milochondrial carrier proteins		56182575, 22278995, 56994075, 22278996, 222789975, 22278996, 222789975, 22278996, 222789975, 22278996, 222789975, 22278996, 222789975, 22278997, 22278997, 22278997, 22278997, 22278997, 22278997, 224996, 264908, 66712502, 264909, 264905, 264906, 264908, 66712502, 264909, 264905, 264906, 264907, 264512, 265008, 265010, 264011, 264017, 264011,
2548	88179079 (5095, 5096)			UNCLASSIFIED	264486 264488 18108394, 52546365, 22278994, 2564262, 29331822, 22278999, 2254259, 29331825, 29331826, 60432289, 29331827, 29331825, 2933182435, 264511, 265007, 264512, 60433356, 87168559, 264684, 264369, 52644229, 265021, 33657023, 264692, 18108374, 52644332, 264557, 18108380, 18108381, 18108382, 18108384, 18108385, 60432113, 22279000, 222799002, 264563,
2549		94196893 (5097, 5098) Novel Protein sim. GBank gij728837 sp p39194 ALU7_HUMAN - !!!! ALU SUBFAMILY LIM domain containing proteins SQ WARNING ENTRY !!!!	Contains protein domain (PF00412) - struct LIM domain containing proteins	struct	56182575, 22278996, 22278997, 22278998, 22278999, 264259, 264508, 264908, 28331830, 285009, 265010, 265018, 264688, 21906764, 21906766, 21906766, 21906769, 265021, 52644150, 264691, 18108368, 60431602, 18108376, 3650643, 3618232, 31908387, 264567
2550	87778584 (5099, 5100)	87778584 (5099, 5100) Novel Protein sim. GBank gil2143886 pir 152523 - nucleoporin p62 homolog - rat (fragment)		UNCLASSIFIED	56182575, 29331822, 264105, 264512, 18108351, 35695917, 264637, 264638

18108396, 65274572, 22278995, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 264905, 56182435, 265007, 60433438, 55812038, 21906754, 65274444, 265017, 265018, 264605, 265019, 264288, 21906766, 21906769, 256220, 60170615, 224693, 33657109, 35666423, 264638, 516182323, 83373044, 22279000	56182575, 35696286, 29331824, 29331826, 35696052, 29331828, 264508, 264907, 56182435, 265008, 264591, 33109954, 264760, 55811957, 35695917, 33657023, 33657109, 18108374, 55811576, 35696423, 35695855, 56182323, 264558	264686, 264488, 263976, 264768, 29331826, 35696052, 35696423, 264601, 264511, 264602, 264910, 264634, 264760, 264555, 264762, 264906, 264592, 264691, 264566, 264908, 264684, 264567, 264909, 264766	22278997, 29331822, 264508, 21906769, 33657023, 33657109, 56182323	29331824, 263972	22278998, 264509, 33657402, 264683, 264684, 264766, 264689, 33657023, 33657109, 35695855, 264558, 264567	264595	22278997, 29331824, 66714117, 29331825, 264906, 264511, 265018, 264448	22278997, 29331822, 29331828, 60433356, 265011, 264288, 264765, 264766, 264769, 21906765, 21906766, 60432113, 264482		22278995, 35696286, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 265007, 60432229, 87168559, 265017, 265018, 265019, 264689, 21906766, 21906769, 35695917, 265020, 33657023, 33657109, 18108374, 264634, 264559, 18108385, 87168518, 22279002
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	cadherin	nuclease	UNCLASSIFIED		UNCLASSIFIED	UNCLASSIFIED		UNCLASSIFIED
Contains protein domain (PF00561) - UNCLASSIFIED alpha/beta hydrolase fold					Contains protein domain (PF00514) - UNCLASSIFIED Armadillo/beta-catenin-like repeats				Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	
95308400 (5101, 5102) Novel Protein sim. GBank gi 4337103 gb AAD18079 - (AF129756) NG26 [Homo sapiens]	(4)	95308243 (5105, 5106) Novel Protein sim. GBank gi 1711658 sp P54797 110_MOUSE - SER/THR-RICH PROTEIN T10 IN DGCR REGION	87761520 (5107, 5108) Novel Protein sim. GBank gi 728835 sp P39192 ALU5_HUMAN - !!!! ALU SUBFAMILY SC WARNING ENTRY !!!!	87627551 (5109, 5110) Novel Protein sim. GBank gi 4884319 emb CAB43260.1 - (AL050084) hypothetical protein [Homo sapiens]	87645533 (5111, 5112) Novel Protein sim. GBank gi 4106984 (AC003038) - R30923_1 [Homo sapiens]	(4)	87617591 (5115, 5116) Novel Protein sim. GBank gi 119110 sp P03211 EBN1_EBV - EBNA-1 NUCLEAR PROTEIN	88096382 (5117, 5118) Novel Protein sim. GBank gi 4538998 emb CAB39619.1 - (AL049481) AlG1-like protein [Arabidopsis thaliana]	87994530 (5119, 5120) Novel Protein sim. GBank gi 5051399 emb CAB44995.1 - (ALD78630) 573K1.3 (mm17M1-4 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor LIKE) protein)) [Mus musculus]	88176575 (5121, 5122) Novel Protein sim. GBank gil5326825gblAAD42056.1AF04495 - (AF044953) NADH:ubiquinone oxldoreductase PGIV subunit [Homo sapiens]
	2552 95332620 (5103, 5104)					2557 79437803 (5113, 511				2561 88176575 (5121, 512

	,	,	_		_	,
56994075, 22278996, 22278997, 22278999, 264259, 29331822, 6043289, 33657402, 60433356, 21906765, 55811957, 60170615, 33657023, 264693, 35695855, 87168518	264488, 35696286, 22278999, 264259, 29331822, 29331824, 35696052, 264508, 264907, 264908, 264909, 52640445, 264510, 264511, 265009, 264910, 264591, 265019, 265018, 265019, 18108351, 264686, 21906767, 21906768, 264691, 264691, 2564693, 27486262, 264628, 18108374, 35696423, 264632, 264634, 264634, 264634, 264634, 264634, 264655, 264568, 18108384, 87168518, 22279000, 22279002, 264482, 264563, 264566, 264486	29331822, 265007, 265010, 265019, 264769, 55811576, 56182323	66714117, 264909, 263978, 264632	18108370, 35695855, 264556, 264558, 18108383	265020, 60170615	60424179, 18108334, 56181686, 56994075, 22278999, 264490, 264259, 29331822, 56182181, 29331824, 60424269, 29331822, 29331826, 29331826, 29331826, 29331826, 29331826, 29331826, 29331826, 29331826, 29331826, 265007, 265009, 264448, 264683, 264288, 18108354, 264769, 21906766, 21906767, 3565917, 265021, 33657023, 18108362, 33657109, 33657
UNCLASSIFIED	UNCLASSIFIED	sulfotransferase	UNCLASSIFIED	struct		synthase
				Contains protein domain (PF00063) - Myosin head (motor domain)		
2562 87645539 (5123, 5124) Novel Protein sim. GBank gild 106984 (AC003038) - R30923_1 [Homo sapiens]	88095497 (5125, 5126) Novel Protein sim. GBank gi∤4866447lemb CAB43371.1 - (AL050270) hypothetical protein [Homo sapiens]	80502783 (5127, 5128) Novel Protein sim. GBank gil1352944 sp P47179 YJ9P_YEAST - HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DAL5 INTERGENIC REGION PRECURSOR		80224956 (5131, 5132) Novel Protein sim. GBank gi[628012 pir A53933 - myosin 1 Contains protein domain (PF00063) - struct Myosin head (motor domain)		91233099 (5135, 5136) Novei Protein sim. GBank gil466009 sp P34548 YNJ4_CAEEL - HYPOTHETICAL 33.8 KD PROTEIN R10E11.4 IN CHROMOSOME III
87645539 (5123, 5124)	88095497 (5125, 5126)		85530906 (5129, 5130)	80224956 (5131, 5132)	86143590 (5133, 5134)	91233099 (5135, 5136)
2562	2563		2565		2567	

	, , ,		
18108394, 56182575, 56181686, 22278995, 22278996, 56994075, 35696286, 22278997, 22278998, 264259, 60432049, 22278998, 22278999, 264259, 60432049, 29331822, 26331824, 29331825, 5669622, 264905, 264906, 56182435, 265008, 265009, 264910, 6043229, 264592, 60433356, 60433438, 21906764, 87168559, 265017, 265018, 265019, 264682, 21906765, 21906766, 21906767, 21906768, 29148627, 21906769, 264691, 264692, 264692, 264692, 264692, 264692, 264692, 264692, 264692, 264693, 35695763, 20281069, 263974, 18108381, 56182323, 18108382, 35526486, 87168318, 22279000, 22279000, 224566	22278996, 29331822, 29331824, 66714117, 29331825, 60432289, 29331827, 35696052, 264907, 264510, 265007, 265009, 264758, 33109954, 265019, 264686, 264689, 265020, 265021, 264691, 264692, 264693, 18108374, 35695855, 264634, 56182323, 264639, 60170394, 83373044, 22279002, 264482	264508, 264905, 264907, 264828, 18108351, 264555, 264556, 264557, 264558, 264559	35696286, 29331827, 35696052, 264100, 264104, 264104, 264110, 264592, 21906754, 29148627, 29148629, 263972, 263974, 18108374, 263976, 35695855, 60170394, 264559, 18108385
transcriptfactor	phosphatase	struct	UNCLASSIFIED
Contains protein domain (PF00170) - Iranscriptfactor	Contains protein domain (PF01240) - phosphatase Protein phosphatase 2A regulatory subunit PR55		
2569 95313764 (5137, 5138) Novel Protein sim. GBank gilg599560 gb AAB84166.1 - (AF029674) basic leucine zipper protein LZIP [Homo sapiens]	94136754 (5139, 5140) Novel Protein sim. GBank gil4758954[ref NP_004567.1 pPPP2 - protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), beta isoform	87733750 (5141, 5142) Novel Protein sim. GBank gij732218[splP34609]YO60_CAEEL - HYPOTHETICAL 128.6 KD PROTEIN ZK1098.10 IN CHROMOSOME III	2572 87627560 (5143, 5144) Novel Protein sim. GBank gil4884319lemb CAB43260.1 - (ALD50084) hypothetical protein [Homo sapiens]
95313764 (5137, 5138).	136754 (5139, 5140)	7733750 (5141, 5142)	7627560 (5143, 5144)
	2570 94	2571 8	80"

2573	95313929 (5145, 5146)	95313929 (5145, 5146) Novel Protein sim. GBank	Contains protein domain (PF00386) - complement	complement	264488, 60424179, 65274572, 56182575,
		gij399138 sp[P02745 C1QA_HUMAN - COMPLEMENT C1Q C1q domain	C1q domain		56181686, 22278995, 56994075, 22278997,
		SUBCOMPONENT, A CHAIN PRECURSOR			60432049, 264259, 29331822, 29331824,
					29331825, 60432289, 29331826, 29331827,
					29331828, 264104, 264107, 264508, 264906,
					29331830, 264909, 264510, 265006, 264512,
					265008, 265009, 264910, 264591, 264592,
					60432229, 264593, 60433356, 264594,
					60433438, 264595, 55812038, 264759,
					21906754, 87168474, 265010, 265011,
					87168559, 265017, 265018, 265019, 264761,
					264762, 264763, 264764, 264369, 264288,
					264685, 264766, 264686, 264687, 264688,
					264769, 56181562, 264689, 21906765,
					21906766, 21906767, 29148627, 21906768,
	_				21906769, 265020, 265021, 265022,
					60170615, 264690, 52644150, 264691,
					264692, 33657023, 65274620, 18108365,
					18108368, 27486265, 60431602, 264629,
					60431528, 263976, 65274791, 35695855,
					20281071, 60431850, 264637, 264638.
	_				264558, 264639, 56182323, 60170394.
					83373044, 18108384, 87168518, 60432113,
					264482, 264564, 264565, 264566, 264567
2574	94746814 (5147, 5148) Novel Protein sim. G	Novel Protein sim. GBank gi 3334982 (AC005306) -	Contains protein domain (PF00651) - UNCLASSIFIED	UNCLASSIFIED	22278995, 264259, 60432289, 29331827,
		R27216_1 [Homo sapiens]	BTB/POZ domain		29331828, 33656970, 264908, 265008,
	_				264910, 264591, 33657402, 265018, 265019,
					264448, 264764, 264369, 264288, 18108357,
					21906765, 21906766, 21906768, 55811957,
					60170615, 264691, 33657023, 264693,
	_				33657109, 33657182, 27486261, 27486264,
					33657349, 264636, 264555, 83373044,
3636		Maria in Commission of the Com			10 100303, 204402 202204 202204 202002 202000
0,07		d / 24406 (3149, 3150) Novel Protein sim. CBank		UNCLASSIFIED	264910, 264601, 264663, 264669, 264060
		gij4929/29[gi]AAD34 123. 1 AF 13100 - (AF 131000) CGF 130 protein [Homo canions]	5		
2576	95357881 (5151, 5152)	95357881 (5151, 5152) Novel Protein sim. GBank	Contains protein domain (PF00442) - ubiquitin	ubianitin	264259, 35696052, 264906, 60433438.
		gil46806811gblAAD27730.1JAF13295 - (AF132955) CGI-21			264681, 18108351, 264288, 52644150,
		protein [Homo sapiens]			264628, 35696423
2577		86996621 (5153, 5154) Novel Protein sim. GBank gil4337103lgb AAD18079 -		UNCLASSIFIED	29331825, 265018, 265019, 264685
		(AF129756) NG26 [Homo sapiens]			
2578	87786941 (5155, 5156)			UNCLASSIFIED	264488, 264906, 264908, 264910, 264596,
					[264603, 264604, 264605, 264766, 21906769, [254578, 25457, 25457]
25.70	87202870 (6157 6159)				204020, 204030, 204030, 204030, 204030
6202	10/29/20/9 (313), 3130)		1	UNCLASSIFIED	29331822, 29331824, 264767

								_			
265007, 265018, 264762	56994075, 29331824, 29331826, 29331828, 264905, 60433356, 60433438, 264758, 87168559, 21906769, 265022, 35695855, 263981	264488, 264907, 264908, 264910, 264764, 264684, 264766, 264638, 264555, 264565	56181686, 264259, 264510, 264512, 264591, 264592, 264593, 264594, 264595, 264596, 264603, 264629, 55810764, 264630, 264637, 264565	264908, 264910, 264768, 264693, 18108374, 55811576, 56182323	264768	22278998, 264259, 29331822, 29331824, 29331827, 29331828, 264906, 265007, 265009, 264591, 60433356, 33657402, 265018, 264762, 264288, 21906766, 21906769, 265222, 264691, 83373044, 56526486, 22279002	264908, 265019, 264768, 264693, 55811576, 56182323	264564	35696052, 264905, 264906, 264907, 264908, 264909, 265009, 265018, 264769, 35696423, 264636	22278995, 29331830, 265008, 265010, 265017, 264639	264259, 29331822, 264106, 264906, 56182435, 265007, 265008, 33109954, 264448, 55811957, 265020, 18108370, 55811576, 22279002
	collagen	eph	transcriptfactor	UNCLASSIFIED	kinase			UNCLASSIFIED	UNCLASSIFIED	glucoamylase	eph
	Contains protein domain (PF00595) - collagen PDZ domain (Also known as DHR or GLGF).		Contains protein domain (PF00047) - transcriptfactor Immunoglobulin domain							Contains protein domain (PF01055) - glucoamylase Glycosyl hydrolases family 31	
2580 88166788 (5159, 5160) Novel Protein sim. GBank gi[2588628 (AC003080) - Similar to AB002297 (PID:g2224539) [Homo sapiens]	87899048 (5161, 5162) Novel Protein sim. GBank gil4406642 gb AAD20049 - (AF131809) Unknown [Homo sapiens]	87786789 (5163, 5164) Novel Protein sim. GBank gi[2739367 (AC002505) - putative phosphatidyllnositol-4-phosphate 5-kinase [Arabidopsis [thaliana]	91220950 (5165, 5166) Novel Protein sim. GBank gil4379112 emb CAA16521.11 - Contains protein domain (AL021578) dJ453C12.2 (similar to transcription factor RBP- Immunoglobulin domain L) [Homo sapiens]		80436126 (5169, 5170) Novel Protein sim. GBank gil2736151 (AF021935) - mytonic dystrophy kinase-related Cdc42-binding kinase [Rattus norvegicus]				85515607 (5177, 5178) Novel Protein sim. GBank gi 3021598 emb GAA71415 - (Y10389) nuclear protein [Xenopus laevis]	87054526 (5179, 5180) Novel Protein sim. GBank gij2104689 (U92793) - alpha glucosidase II, alpha subunit [Mus musculus]	94192167 (5181, 5182) Novel Protein sim. GBank gil5702202lgb AAD47199.1 AF12916 - (AF129166) long- chain acyl-CoA synthetase 5 [Homo sapiens]
88166788 (5159, 5160)	87899048 (5161, 5162)	87786789 (5163, 5164)	91220950 (5165, 5166)	80430941 (5167, 5168)	80436126 (5169, 5170)	91226136 (5171, 5172)	80430943 (5173, 5174)	80074385 (5175, 5176)	85515607 (5177, 5178)	87054526 (5179, 5180)	94192167 (5181, 5182)
2580	2581	2582	2583	2584	2585	2586		2588	2589		2591

68		2.	2,				
18108397, 56182575, 35696286, 56994075, 264259, 29331822, 29331824, 29331828, 2643289, 29331827, 29331828, 2643289, 29331827, 29331828, 264906, 264909, 265007, 265008, 264910, 60433229, 264594, 60433256, 60433438, 55812038, 1108348, 21906754, 265011, 87168559, 265017, 265019, 264764, 264369, 264288, 264766, 265021, 60170615, 33657023, 33657109, 264629, 35699423, 35695855, 264557, 264638, 60170394, 56182323, 264482, 264466, 87168518, 264563, 264482, 264466	22278999, 29331825, 264758, 21906754, 52646317, 265010, 18108351, 264288, 264369, 21906768, 264693, 18108370, 264637, 264638, 264482	264486, 18108398, 56182575, 35696286, 22278997, 264093, 264259, 29331822, 29331822, 26331822, 2643045, 56182435, 264906, 2644046, 56182435, 264510, 264512, 265007, 264757, 21906754, 87168474, 255017, 264760, 264448, 264766, 26489, 21906768, 33657109, 263975, 263977, 264634, 264566, 264566, 264566, 264567, 264566, 264567, 264	22278999, 264259, 265018, 264448, 265021,	60431528 264905, 264509, 264908, 264762, 264765, 35695917, 35695855, 264635, 264636, 83373044, 264486	22278998, 264259, 29331824, 87168474, 264683, 21906766, 35695917, 264691, 33657023, 33657109, 18108370, 18108374, 264564, 264565	264766, 264769, 21906768, 33657182, 35695763, 18108370, 18108374, 264635, 264636, 56526486, 22279000, 264566	35696286, 264093, 264288, 21906769, 35696423, 35695855
transport	tm7	UNCLASSIFIED	UNCLASSIFIED UNCLASSIFIED	UNCLASSIFIED	ÜNCLASSIFIED	MHC	kinasereceptor
							Contains protein domain (PF00400) - kinasereceptor WD domain, G-beta repeat
Novel Protein sim. GBank gil3024998 sp Q60936 YAB1_MOUSE - HYPOTHETICAL HEART PROTEIN	87754416 (5185, 5186) Novel Protein sim. GBank gil4929729 gb AAD34125.1 AF15188 - (AF151888) CGI-130 protein [Homo sapiens]	95305758 (5187, 5188) Novel Protein sim. GBank gil4929587[gb AAD34054.1 AF15181 - (AF151817) CGI-59 protein [Homo sapiens]	79561676 (5189, 5190) 87538537 (5191, 5192) Novel Protein sim. GBank gild309681 gb AAD15478 -	(AC006930) R33423_1 [Homo sapiens]	88094948 (5195, 5196) Novel Protein sim. GBank gi[1001351[dbj]BAA10838] - (D64006) hypothetical protein [Synechocystis sp.]	87842889 (5197, 5198), Novel Protein sim. GBank gij3941737 (AF109719) - BAT2 [Mus musculus]	2600 87787846 (5199, 5200) Novel Protein sim. GBank gil4263521lgb AAD15347 - (AC004044) putative WD-repeat protein [Arabidopsis thaliana]
95332648 (5183, 5184)	87754416 (5185, 5186)	95305758 (5187, 5188)		94784089 (5193, 5194)			87787846 (5199, 5200)
2592	2593	2594	2595 2596	2597	2598	2599	2600

9 4406632 gb AAD20047 - mo sapiens UNCLASSIFIED Oncogene B 1 PRRP2 - RAS-related on Oryctolagus cuniculus Oryctolagus Orycto	91243070 (5201, 5202)	2601 [91243070 (5201, 5202) Novel Protein sim. GBank	kinase	56182575, 22278999, 264259, 29331822,
### ### ##############################		gij728837jspjP39194jALU7_HUMAN - II!! ALU SUBFAMILY SO MABNIMG ENTEX III	,	29331824, 60432289, 29331826, 29331827, 264906, 265007, 265008, 264591, 60433356
### ##################################				33657402, 60433438, 21906754, 265011,
### MOUSE - LIGATIN ### MOUSE - LIGATIN ### MOUSE - LIGATIN #### MOUSE - LIGATIN #### MOUSE - LIGATIN ###################################				265018, 265019, 18108351, 264448, 264369,
### ### ##############################				21906769, 265020, 60170615, 264693,
### MOUSE - LIGATIN WINCLASSIFIED UNCLASSIFIED Oncogene UNCLASSIFIED Oncogene				33657109, 18108370, 18108376, 56182323,
### ### ##############################				18108381, 18108385, 22279002, 264563
A_MOUSE - LIGATIN In price - Ligatin and a concodent and a concodent and a concodent all and a confidence and a confidence a confidence and a confidence and a confidence and a confidence a confidence and a con	80022 (5203, 5204)	Novel Protein sim. GBank gi 4406632 gb AAD20047 -		60433438, 21906754, 87168559, 264601,
A_MOUSE - LIGATIN UNCLASSIFIED oncogene i.1pRRP2 - RAS-related on gij3628745 dbj BAA33366 - 3 [Oryctolagus cuniculus]		(AF131801) Unknown [Homo sapiens]		264369, 264288, 21906767
ancogene	25821 (5205, 5206)	Novel Protein sim. GBank	UNCLASSIFIED	264488, 65274572, 22278995, 22278996,
oncogene		gij3122367jsp Q61211 LIGA_MOUSE - LIGATIN		56994075, 22278997, 22278998, 22278999,
oncogene				60432049, 264259, 29331822, 29331824,
oncogene UNCLASSIFIED				60432289, 29331826, 29331827, 29331828,
oncogene				35696052, 56182435, 264113, 265008,
oncogene UNCLASSIFIED				265009, 60433356, 264757, 60433438,
oncogene				264759, 33657084, 8/1684/4, 265010,
oncogene UNCLASSIFIED				265011, 87168559, 265017, 265018, 265019,
oncogene ONCLASSIFIED				264448, 264683, 18108354, 264288, 264767,
oncogene UNCLASSIFIED				264689, 21906765, 21906766, 21906767,
oncogene UNCLASSIFIED				21906768, 21906769, 55811957, 265020,
ancogene				265021, 265022, 60170615, 264691,
oncogene UNCLASSIFIED				33657023, 264693, 33657109, 27486262,
oncogene UNCLASSIFIED				18108374, 35696423, 65274791, 35695855,
oncogene UNCLASSIFIED				264555, 264636, 264637, 56182323,
oncogene UNCLASSIFIED				83373044, 56526486, 87168518, 60432113,
Oncogene				22279000
UNCLASSIFIED	76601 (5207, 5208)	Novel Protein sim. GBank	ancogene	264259, 35696052, 264508, 264906, 264907,
UNCLASSIFIED		gij5454030jrefjNP_006468.1jpRRP2 - RAS-related on		264908, 264909, 264510, 264512, 265008,
UNCLASSIFIED		chromsome 22		264910, 33657402, 264604, 264605, 264762,
UNCLASSIFIED				264763, 264682, 264764, 264683, 264768,
UNCLASSIFIED				264769, 264689, 33657023, 264693,
UNCLASSIFIED				18108365, 264628, 35696423, 264631,
UNCLASSIFIED				264632, 264634, 264635, 264637, 18108381,
UNCLASSIFIED				264639, 83373044, 264565
(AB013721) mitsugumin 23 [Oryctolagus cuniculus]	116756 (5209, 5210)	Novel Protein sim. GBank gij3628745 dbj BAA33366 -	UNCLASSIFIED	22278998, 264490, 60432049, 264259,
		(AB013/21) mitsugumin 23 [Oryctolagus cuniculus]		60432289, 264909, 265006, 60433336, 60433438 964758 91606754 965010
				00453430, 204730, 21300734, 203010,
				263011, 203016, 203061, 19106391, 203260, 264766, 264685, 21906765, 21906766,
				21906768, 21906769, 264691, 264692,
		-		264693, 65274791, 264634, 264555, 264636
122279002. 284566	46406 (5211, 5212)			22278996, 264510, 264512, 265009, 264766,
				22279002, 264566

29331822, 29331825, 29331826, 29331827, 29331828, 264907, 264908, 265712502, 264907, 264907, 264908, 265712502, 264828, 56182435, 55812038, 265010, 265017, 265018, 265019, 264689, 21906765, 55811957, 265020, 265022, 264692, 33657103, 264639, 56182323, 25457109, 18108370, 264639, 56182323, 254509, 264907, 264639, 264567, 264567	264631	264488, 65274572, 35696286, 22276997, 22278999, 60432049, 264259, 56182181, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 264905, 264907, 80433356, 60433438, 55812038, 285011, 874288, 264765, 264689, 21906765, 21906767, 21906769, 265020, 265021, 264691, 264692, 33657109, 27486561, 18108370, 65274791, 264636, 264556, 56182323, 18108385, 56526486	18108394, 22278994, 56994075, 60432049, 264229, 284259, 29331822, 29331825, 60432289, 29331827, 264107, 264109, 264905, 26182435, 264112, 265006, 265007, 265009, 265009, 6043356, 60433438, 265011, 26428, 26428, 265017, 26448, 26488, 26507, 33657023, 263976, 55811576, 264538, 87168518, 264638, 87168518, 2645113	22278999, 265017, 264684, 21906768, 22279000	265009, 264910, 264759, 265017, 21906767, 18108365, 18108388, 60432113	264685	22278998, 22278999, 29331825, 264508, 264906, 21906754, 264602, 264766, 264769, 52644229, 21906765, 33657109, 27486264, 18108370, 263972, 264555, 60432113	264768, 18108394, 264692, 264693, 264508, 264509, 264907, 264602, 264908, 264909, 18108377, 264511, 264512, 264910, 264635, 264595, 265010, 264404, 264563, 264764, 264585, 264766
		histone	transcriptfactor	isomerase	transport	UNCLASSIFIED		UNCLASSIFIED
		Contains protein domain (PF00850) - histone Histone deacetylase family		Contains protein domain (PF00254) - isomerase FKBP-type peptidyl-prolyl cis-trans isomerases			Contains protein domain (PF00625) - Guanylate kinase	
87627742 (5213, 5214) Novel Protein sim. GBank gild826626[gb]AAD30202.1] - (AF135022) mediator [Homo sapiens] (AF135022) mediator [Homo sapiens] (AF135022) mediator [Homo sapiens]	function unknown [Homo sapiens]	94843791 (5217, 5218) Novel Protein sim. GBank gil3024889 splP56524 Y288_HUMAN - HYPOTHETICAL PROTEIN KIAA0288 (HA6116)	88177654 (5219, 5220) Novel Protein sim. GBank gil433685[gb AAD17969] - (AF106473) leucine-rich-domain inter-acting protein 1; LER inter-acting protein 1; LEAP1 [Mus musculus]	87428890 (5221, 5222) Novel Protein sim. GBank gij3876761 emb CAA92994 - Contains pro (268760) predicted using Genefinder: Similarity to Mouse FKBP-type protein (SW:FKB3_MOUSE) [Caenorhabditis Isomerases elegans]	87771198 (5223, 5224) Novel Protein sim. GBank gi[5679136]gb AAD46874.1 AF16093 - (AF160934) BcDNA.LD14189 [Drosophila melanogaster]		87643948 (5227, 5228) Novel Protein sim. GBank gi[5533081[gb AAD45009.1 AF16118 - (AF161181) P55T protein [Mus musculus]	
				87428890 (5221, 5222	87771198 (5223, 5224	3 79481496 (5225, 5226)	87643948 (5227, 5228	87381996 (5229, 5230)
2607	3	2609	2610	2611	2612	2613	261	2615

	576,	3.		22 8, 88, 89, 87, 88, 89, 87, 87, 87, 87, 87, 87, 87, 87, 87, 87		336.	
22278995, 22278997, 22278998, 60432049, 60432298, 26432249, 60433456, 264594, 60433438, 33109954, 87168474, 265011, 265017, 265019, 264288, 264766, 21906765, 21906767, 265020, 265021, 18108376, 18108377, 18108387, 87168518, 264482, 264567	265010, 265019, 264369, 264693, 55811576, 22279002	264489, 22278996, 264490, 264259, 29331822, 264102, 264509, 264906, 264906, 265009, 66712502, 29331830, 265008, 264910, 265009, 60433356, 60433438, 264758, 21906754, 265011, 87168559, 265017, 265018, 264689, 21906765, 21906766, 21906766, 21906765, 26502, 33657023, 264689, 23657109, 264628, 18108374, 35695855, 18108381, 83373044, 18108385, 18108381, 264568, 21966786, 21966787, 26505488, 264628, 18108385, 264628, 264588, 264628,	264259, 60432289, 60433438, 21906754, 264369, 60432113, 264566	52644507, 52645156, 52646842, 65274572, 22278995, 56994075, 35696286, 22278999, 60432049, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331822, 29331824, 28331825, 29331826, 265008, 3669022, 265008, 6043336, 365010, 265017, 265018, 264448, 264369, 264689, 21906764, 2650817, 2654768, 5264429, 264688, 264689, 21906765, 21906768, 35695917, 52644150, 264622, 33657109, 3569587, 35696423, 264565, 52644332, 18108382, 83373044, 18108385, 18108387, 87168518, 60432113, 222279002	264594, 264636	264488, 264906, 264907, 264908, 264512, 265007, 264758, 35695917, 264634, 264636, 264563, 264482	29331824, 35696052, 265007, 265010, 264288, 29148629
isomerase	kinase	UNCLASSIFIED	dehydrogenase	UNCLASSIFIED			UNCLASSIFIED
Contains protein domain (PF00254) - isomerase FKBP-lype peptidyl-protyl cis-trans isomerases		Contains protein domain (PF00397) - UNCLASSIFIED WW domain	Contains protein domain (PF00465) - dehydrogenase Iron-containing alcohol dehydrogenases	Contains protein domain (PF00628) - UNCLASSIFIED PHD-finger			
gij3876761femb CAA929941 - Senefinder; Similarity to Mouse v:FKB3_MOUSE) [Caenorhabditis	86976888 (5233, 5234) Novel Protein sim. GBank gij728831 sp P39188 ALU1_HUMAN - !!!! ALU SUBFAMILY J WARNING ENTRY !!!!	91231662 (5235, 5236) Novel Protein sim. GBank gij3319282 (AF049103) - Huntingtin interacting protein [Homo sapiens]	87694000 (5237, 5238) Novel Protein sim. GBank gi[2431772 (U66411) - putative type III alcohol dehydrogenase [Drosophila melanogaster]	95314841 (5239, 5240) Novel Protein sim. GBank gil4322567lgb AAD16097 - (AF090436) dachshund variant 1 [Mus musculus]		81780390 (5243, 5244) Novel Protein sim. GBank gil4557341 ref NP_001174.1 pATP6 - ATPase, H+ fransporting, lysosomal subunit 1; vacuolar proton pump; H- ATPase subunit	91639306 (5245, 5246) Novel Protein sim. GBank gij3880355[emb]CAB05299] - (Z82285) predicted using Genefinder [Caenorhabditis elegans]
87428895 (5231, 5232)	86976888 (5233, 5234)	91231662 (5235, 5236)	87694000 (5237, 5238)	95314841 (5239, 5240)	80253495 (5241, 5242)	81780390 (5243, 5244)	91639306 (5245, 5246)
2616		2618	2619		2621		2623

624	91639308 (5247, 5248)	2624 91639308 (5247, 5248) Novel Protein sim. GBank gij3880355[emb CAB05299] - (Z82285) predicted using Genefinder [Caenorhabditis elegans]		UNCLASSIFIED	56181686, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 56182181, 29331825, 60432289, 29331828, 35696052, 29146499, 66712502, 52644045, 265007, 265008, 6043335, 33109954, 264288, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 35695855, 18108337, 60431528, 55811576, 35695855, 56182323, 18108386, 87168518, 22279002, 18108391	
	86452068 (5249, 5250)	86452058 (5249, 5250) Novel Protein sim. GBank gij2887429jdbjjBAA24857 - (AB007887) KIAA0427 [Homo sapiens]		UNCLASSIFIED	264091, 264511, 263981	
_	16533797 (5251, 5252)	16533797 (5251, 5252) Novel Protein sim. GBank gil487416 (L20302) - actin filament protein [Gallus gallus]		struct	265008	
	87636823 (5253, 5254)	87636823 (5253, 5254) Novel Protein sim. GBank gij88462 pir A27307 - proline-rich phosphoprotein (gene PRH1, Db allele) - human		UNCLASSIFIED	22278996, 265007, 265009, 264448, 21906767, 265021, 264558, 87168518	
	94848254 (5255, 5256)	94848254 (5255, 5256) Novel Protein sim. GBank gi[3123552 emb CAA18609 - (AL022578) d.1393P12.2 (hypothetical Proline-rich protein KIAA0269 LIKE) (Homo sapiens)		UNCLASSIFIED	22278997, 22278999, 264259, 60432049, 29331822, 29331824, 29331825, 29331825, 29331827, 35696052, 29331828, 264907, 264909, 265008, 264591, 60433356, 60433438, 265010, 265017, 265018, 264369, 264288, 18108357, 21906765, 21906768, 265022, 622790791, 264638, 18108387, 87168518, 22279079	
	87376490 (5257, 5258)	87376490 (5257, 5258) Novel Protein sim. GBank gjl4929595[gb AAD34058.1 AF15182 - (AF151821) CGI-63 protein [Homo sapiens]		synthase	29331826, 29331826, 264102, 265006, 264766, 35695917, 264691, 33657023, 263972, 18108374, 22279000	·
	79188364 (5259, 5260)			UNCLASSIFIED	264636, 18108385	_
	94845909 (5261, 5262)	Novel Protein sim. GBank gij321605 prt JQ1161 - Gag protein - Visna virus (strain EV1)	Contains protein domain (PF00098) - dna_ma_bind	dna_ma_bind	52644507, 52645156, 52646365, 52646842, 22278994, 22278995, 35696286, 56994075, 22278994, 22278995, 35696286, 56994075, 22278997, 22278999, 264259, 25645080, 29331822, 29331822, 29331822, 29331822, 29331822, 29331822, 29331822, 26431826, 264908, 264311, 264512, 265007, 265008, 264910, 5264511, 265017, 265018, 264010, 87168559, 265017, 265018, 264010, 87168559, 264448, 264288, 264369, 264760, 264762, 2564429, 21906765, 21906766, 21906767, 21906767, 21906767, 21906767, 21906767, 23695917, 265020, 32657182, 32645129, 33657193, 33657183, 3569643, 264486, 264486, 264486	
2632	36730414 (5263, 5264)				264685	_

22278995, 22278997, 22278998, 264259, 29331822, 29331824, 29331825, 29331827, 264508, 264508, 265008, 265008, 265008, 265009, 56512038, 33657084, 55811386, 265010, 265011, 87168559, 265018, 264683, 264288, 264686, 29148629, 33657023, 264639, 33657182, 35695763, 55811576, 264639, 56182323, 83373044, 18108385, 5652486, 87168518, 22279000, 22279002, 264565	29331826, 263972, 264089	22278997, 264259, 29331824, 66714117, 29331827, 264907, 33657084, 265017, 265018, 264448, 264288, 21906767, 21906768, 29148629, 18108376, 55811576, 35695855, 87168518, 22279000	264569, 29331822, 29331828, 265006, 60170831, 264681, 264765, 264685, 29148627, 21906769, 29148784, 265022, 60170615, 264635, 18108385, 56526486, 22279002, 264567	264488, 264489, 52644507, 264887, 5264585, 52646842, 22278994, 22278996, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 2264508, 66714117, 29331825, 29331825, 29331825, 29331825, 264509, 264509, 264501, 264501, 264501, 264591, 264591, 264591, 264592, 60433256, 264591, 264594, 33657402, 60433356, 264501, 264601, 264602, 264603, 264603, 264603, 264604, 264764, 264761, 264764, 264766, 26481, 264766, 26481, 264764, 264761, 264276, 21906765, 21906766, 21906766, 21906766, 21906766, 21906766, 21906769, 33695917, 269011, 60170615, 264691, 264634, 366964, 18108374, 35696423, 35696423, 35696433, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264638, 264566, 2645
	UNCLASSIFIED			phosphatase
Contains protein domain (PF00084) - Sushi domain (SCR repeat)			Contains protein domain (PF01546) - Peptidase family M20/M25/M40	Contains protein domain (PF00762) - phosphatase, Catalytic domain
Novel Protein sim. GBank gil 1139548 dbi BAA10889 - (D64009) seizure-related gene product 6 type 2 precursor [Mus musculus]	87330921 (5267, 5268) Novel Protein sim. GBank gi[5441611 emb CAB46854.1 - (AJ388555) hypothetical protein [Canis familians]	86623144 (5269, 5270) Novel Protein sim. GBank gil4680663lgb AAD27721.1 AF13294 - (AF132946) CGI-12 protein [Homo sapiens]	87260534 (5271, 5272) Novel Protein sim. GBank gil3879146[emb CAB07646] - (293386) Similarity to Yeast hypothetical 52.9 KD protein (SW:P43616); cDNA EST EMBL:M89432 comes from this gene; cDNA EST EMBL.D71008 comes from this gene; cDNA EST EMBL.D73578 comes from this gene; cDNA EST EMBL.D73578 comes from this gene; cDNA	95011299 (5273, 5274) Novel Protein sim. GBank gil4758208 ref NP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related)
		86623144 (5269, 5270)	87260534 (5271, 5272)	95011299 (5273, 5274)
263	2634	2635	2636	2637

56182575, 22278996, 35696286, 22278998, 264259, 29331822, 56182181, 29331825, 60424269, 60432289, 35696052, 66712502, 264908, 265007, 55812038, 33109954, 21906754, 33657084, 265019, 264448, 264288, 56181562, 21906765, 21906766, 21906768, 21906768, 21906768, 21906768, 21906769, 365937, 265020, 265021, 52244150, 264693, 33657109, 33657349, 60431528, 18108374, 55810764, 25696423, 56182323, 60432113, 22279902, 264554	29146498, 55182436, 33109954, 265011, 264682, 55811957, 35695917, 264690, 263976, 18108377, 35696423, 60432113	56182575, 35696286, 22278998, 29331824, 29331826, 60432289, 66712502, 56182435, 60170831, 60432229, 33657402, 33109954, 21906754, 265017, 264686, 264688, 21906768, 60170515, 264693, 2539574, 18108370, 263976, 60170394,	264685	29331824, 29331826, 29331827, 265007, 55812038, 21906754, 18108366, 18108384, 22279002, 264567	264692	265018, 18108370, 18108387, 264566	60424179, 65274572, 56182575, 264259, 56182181, 264908, 56182435, 55811957, 35695917, 265021, 263976, 55810764,	03/4791, 3018/343, 833/3044, 652/4727 56182575, 56181686, 264092, 264259, 56182181, 6043259, 264907, 33657402, 56812088, 21906754, 87168559, 265017, 264448, 264369, 264288, 21906765, 21906766, 21906767, 21906768, 33657109, 18108370, 264628, 55811576, 26456,	264693
UNCLASSIFIED		UNCLASSIFIED		synthase	UNCLASSIFIED	UNCLASSIFIED			
				Contains protein domain (PF00054) - synthase Laminin G domain				Contains protein domain (PF00097) - transcriptfactor Zinc finger, C3HC4 type (RING finger)	
			Novel Protein sim. GBank gil5360271 dbj BAA81908.1 - (AB029335) HrPET-3 [Halocynthia roretzi]	Bodoo 45 (5301, 5302) Novel Protein sim. GBank gil4240225[dbj BAA74891.1 - (AB020675) KIAA0868 protein [Homo sapiens]		Novel Protein sim. GBank gil4493956 emb CAB11123.2 - (798551) predicted using hexExon: MAL3P6.28 (PFC0845c), Hypothetical protein, len: 167 aa: Similarity to model organism hypothetical proteins (C.elegans, D.melanogaster, S.cerevisiae & S.pombe). C.elegans protein ZK287.5 (TR:		91229018 (5309, 5310) Novel Protein sim. GBank gij3875272[embjCAB02861] - (Z81051) predicted using Genefinder; similar to Zinc finger. C3HC4 type (RING finger); cDNA EST yk443h5.3 comes from this gene; cDNA EST yk443h5.5 comes from this gene [Caenorhabditis elegans]	84562601 (5311, 5312) Novei Protein sim. GBank gi 3043718 dbj BAA25523 - (AB011169) KIAA0597 protein [Homo sapiens]
	8/60038/ (5295, 5296) 94128783 (5297, 5298) Naviel Drafeis size	(0.00) (0.00) (0.00) (0.00) (0.00)	87297533 (5299, 5300) Novel Protein sim. (AB029335) HrPET	poudd/45 (5301, 5302) 10343426 (6303, 6304)	19343123 (3303, 3304)		95103240 (5307, 5308)	91229018 (5309, 5310)	84562601 (5311, 5312)
2647	2649		2650	2653	2000				2656

	EACC1700 15040 50441	TA COCCA A CITAL COSCACIONA A COCCA A CALLANDA A COCCA A CALLANDA			
) C07	52561726 (5313, 5314)	2230 720 (3313, 3314) Novel Protein sim. Geank gijoobsooy doj BAA63038 (AB029009) KIAA1086 protein [Homo sapiens]			204093
2658	88062454 (5315, 5316)	88062454 (5315, 5316) Novel Protein sim. GBank gij3688089 (AC005757) - R32611_1 [Homo sapiens]	Contains protein domain (PF00560) - nucleaseinhib Leucine Rich Repeat	nucleaseinhib	35696286, 264259, 29331822, 29331824, 29331826, 29331828, 265019, 264683, 21906768, 35695917, 264693, 35695855, 264677, 264697, 264677
2659	87600755 (5317, 5318) Novel Protein sim. (A.1243459) proteot	Novel Protein sim. GBank gi[5420387 emb CAB46679.1 - (A.1243459) proteophosphopican II eishmania maiori	Contains protein domain (PF01426) - UNCLASSIFIED BAH domain	UNCLASSIFIED	254909, 264910, 265018, 264369, 264769, 21906769, 264693, 263972, 18108388
2660	01718472 (5310 5320)	91718472 (5319 5320) Novel Protein cim CRank	Contains protein domain (DE00036) - kinase	pacua	252448 65274572 35606286 22278008
000	911 10412 (3318, 3320)	novei Flotelii sim. GBailk jai7288371spiP39194/ALU7 HUMAN - !!!! ALU SUBFAMILY	Contains protein comain (Process) - EF hand	Alliase	22278999, 264259, 29331822, 29331824,
		SQ WARNING ENTRY IIII			60432289, 29331826, 35696052, 264908.
					56182435, 265008, 265009, 60433356,
					264594, 265010, 265018, 55811150,
					18108351, 264682, 264684, 264369, 264288,
					264687, 21906765, 29148784, 35695917,
					60170615, 52644150, 33657023, 33657109,
					35696423, 35695855, 264556, 60170394,
					18108385, 22279000, 22279002
2661	95342817 (5321, 5322)	95342817 (5321, 5322) Novel Protein sim. GBank		glycoprotein	60432049, 264259, 29331824, 29331825,
		gil4758048 ref NP_004739.1 pCPR8 - cell cycle progression			29331826, 29331827, 29331828, 264906,
		8 protein			264909, 264593, 33109954, 265010, 265017,
					265018, 265019, 264760, 264448, 264369,
					264288, 21906765, 21906768, 265022,
					264691, 33657023, 27486262, 60431528.
					18108374, 35695855, 18108388, 264482
2662					264555, 264556, 264558, 264486
2663	87780623 (5325, 5326)			dehydrogenase	264906, 264909, 264757, 264758, 264767,
		(Z66494) similar to choline dehydrogenase; cDNA EST			264691, 33657023, 264638
		yk346d5.5 comes from this gene; cDNA EST yk346d5.3			
		comes from this gene [Caenorhabditis elegans]			
2664	85518329 (5327, 5328)	85518329 (5327, 5328) Novel Protein sim. GBank gi]1389670 (U58977) - Notch homolog Scatlaned wings II relia supplied	Contains protein domain (PF00008) - oncogene	oncogene	35696286, 264509, 264595, 264288, 264685, 26468
2666		87770669 /6930 6930 Navid Distriction Chart ail 1098/10619		CELECTOR LOIN	26506386 33378000 30331823 36606063
2007		(AL050190) hypothetical protein [Homo sapiens]		ONCLASSIFIED	25090260, ZZZZZ 0393, Z333 10ZZ, 3509003Z, 264906, 264907, 264909, 264510, 264511,
					264512, 264593, 60433438, 265019, 264681,
					21906765, 21906766, 21906767, 21906768,
				-	265020, 265022, 35696423, 35695855,
9990	070004 4003) (5004	1 100			22279002, 264482, 264486
9007	6/6/24/2 (5331, 5332) Novel Protein Sim.	Novel Protein Sim. GBank		UNCLASSIFIED	29331825, 265007, 264910, 60432229,
		[gi 5106956 gb AAD39906.1 AF11361 - (AF113615)			265019, 264288, 21906767, 264558,
1000	1002 00000	FRITTA DOMAIN-CONTRIBING PROTEIN FRUS (Homo sapiens)			222/9002
/997	87422720 (5333, 5334)	8/42Z/ZU (5333, 5334) Novel Protein sim. GBank	Contains protein domain (PF01138) - nuclease	nuclease	264907, 29331830, 264681, 264683, 264288,
		gi 2500570 sp Q17533 RNPH_CAEEL - RIBONUCLEASE bh i ive bbotein bose4 1	3' exoribonuclease family		35695855, 264632, 264556, 264557, 264558, 264660 264600 264660 264600 264660 264600 264600 264600 264660 2646600 2646600 2646600 2646600 2646600 264600 264600 264600 2646
		LITELIA POSO4. I			204000, 204000, 204000, 204001

6 (5335, 5336)		<u> </u>	UNCLASSIFIED	56181686, 35696286, 22278998, 22278999, 56182181, 29331824, 60424269, 29331825,
	gilə+ə+ rool erilivi _cocoət : 1 pz.1 ol. z.1 o.1 oriolog			35696052, 29331828, 66712502, 56182435, 60433356, 264758, 21906754, 55811386,
				265011, 87168559, 265017, 265019, 55811150, 264448, 264369, 264288,
				21906765, 21906766, 21906768, 55811957,
				35595917, 255020, 255021, 33557025, 264692, 33657109, 35695763, 60431528,
				18108374, 35696423, 55811576, 264634,
				60431850, 83373044, 18108385, 87168518,
900	05445794 (2937 2938) Named Bootsin sim CBook ai(2447043)airil (CA800 aradina			222/3000, 264563, 264564 264489, 264689, 21906767, 65274572.
2	-			56182575, 21906768, 29148627, 21906769,
				29148629, 35696286, 35695917, 22278996,
				22278998, 265021, 265022, 60170515, 52544150, 60432040, 264250, 264691
				3264150, 60432043, 204233, 204331, 33657023, 264692, 29331822, 29331824,
				29331825, 60432289, 33657109, 29331826,
				29331827, 35696052, 29331828, 29146498,
				29146499, 264905, 264908, 52644045,
				264909, 56182435, 35696423, 65274791,
				[35695855, 265006, 264910, 264635,
				50432229, 254532, 254636, 35162323,
				00433330, 00170384, 00433430, 2043339,
				100,000,000,000,000,000,000,000,000,000
				60432113, 265019, 264448, 264369, 264684, 264288
5340)	T	Contains protein domain (PF00628) -		18108370, 263974
	gil1723523 sp Q10362 YDBB_SCHPO - HYPOTHETICAL 94.9 KD PROTEIN C22E12.11C IN CHROMOSOME I	PHD-finger		
342	91214936 (5341, 5342) Novel Protein sim. GBank		transport	52646842, 56994075, 264259, 29331822,
				29331824, 29331825, 29331827, 33656970,
	long-chain acyl-CoA synthetase homolog 2; VLCS-H2			264509, 265006, 33109954, 21906754,
	[Homo sapiens]			264682, 264288, 265021, 33657023,
				3355/109,3355/108,2/456251,2/46625. 27485265 18108376 18108385
344	87399123 (5343 5344) Novel Protein sim GBank	Contains protein domain (PF01344) - UNCLASSIFIED	UNCLASSIFIED	264767
:	034677, 11AC00634 - (AC006341)	Kelch motif		
	Contains two PF 01344 Kelch motif domains. Arabidopsis			
i	[thaliana]			

22278995, 22278996, 22278997, 22278999, 29331822, 29331824, 29331828, 35596052, 60433438, 87168474, 87168559, 265017, 265018, 265019, 264448, 21906768, 255020, 265021, 33557109, 27486262, 35695763, 60431850, 60170394, 87168518, 264563	264488, 22278996, 35696286, 264091, 264259, 29331824, 29331826, 35696052, 264511, 55812038, 85658542, 264766, 21906765, 35695917, 264629, 35696423, 18108383, 87168518	264691	264906, 265008	16108394, 65274572, 56182575, 22278994, 22278995, 56984045, 22278996, 35596286, 22278995, 22278999, 264490, 264259, 52645080, 29331824, 29331825, 60432289, 29331837, 29331828, 35698052, 29146499, 29331830, 264908, 255009, 60432229, 60433356, 60433436, 55812038, 265010, 265017, 18108351, 264682, 264448, 264683, 264288, 21906765, 219067	83373044, 18108387, 22279000, 22279002, 264564
polymerase	transcriptfactor	UNCLASSIFIED	UNCLASSIFIED	synthase	
Contains protein domain (PF00644) - polymerase Poly(ADP-ribose) polymerase catalytic region.	Contains protein domain (PF00170) - transcriptfactor bZIP transcription factor			Contains protein domain (PF00501) - synthase AMP-binding enzyme	
Novel Protein sim. GBank gi[5457337 emb CAB41505.2 - (AJ236876) poly(ADP-ribosyl) polymerase-2 [Homo 'sapiens]	Novel Protein sim. GBank gl 4758824 ref NP_004280.1 pNRF3 - nuclear factor (erythroid-derived 2)-like 3		110000000000000000000000000000000000000	obank gillo veokakpirilisez 154 - acetyl- ruit fly (Drosophila melanogaster)	
2673 87430749 (5345, 5346) Novel Protein sim. (AJ236876) poly(Al sapiens]	94847721 (5347, 5348)	79563835 (5349, 5350)	2010 13020333 (3331, 3332)	4500 '5500 DOBENTA	
267	267	2675	287		

	· · · · · · · · · · · · · · · · · · ·	Γ		<u> </u>	Γ
264488, 263994, 264489, 18108394, 52646842, 35696286, 22278999, 264259, 22331825, 35696022, 264508, 264908, 2644509, 264509, 264509, 264509, 264509, 264500, 264510, 265007, 265007, 265007, 265007, 265007, 265007, 264509, 264507, 264509, 264507, 264609, 264601, 2650011, 87168559, 264600, 264601, 264602, 264603, 265017, 264600, 264601, 264602, 264762, 18108351, 264763, 264689, 264762, 264689, 264762, 264689, 21906765, 21906767, 35695917, 265020, 265021, 52644150, 264691, 33657023, 264693, 18108364, 18108365, 33657109, 33657349, 264638, 264637, 264632, 264638, 264638, 264637, 264632, 264638, 264638, 264637, 264632, 264638, 264638, 264637, 264637, 264632, 264638, 26	264488, 22278996, 35696286, 264259, 29331826, 29331827, 29331828, 264909, 56182435, 264113, 264511, 265008, 60433438, 264758, 89568942, 87186474, 265011, 265011, 265019, 264288, 21906764, 21906764, 21906769, 21906769, 258011957, 265021, 264690, 33657023, 55810764, 35696423, 55811576, 264331, 18108381, 60170394, 83373044, 87168518, 264566	264259, 264102, 264905, 264908, 285007, 265008, 33109954, 265011, 18108351, 264768, 33657023, 20281149, 263972, 264630, 264635, 264638	264489, 22278995, 264509, 264905, 264906, 264907, 264907, 264909, 264510, 265006, 264512, 264910, 264594, 60433438, 264758, 264604, 265019, 264605, 264760, 264764, 264687, 264768, 264769, 2906769, 2669517, 265020, 33657023, 264631, 264638, 264638, 264639, 264638, 26468486, 264688, 2646	35696286, 264908, 55811386, 265017, 55811150, 55811957, 35695917, 60431528, 55810764, 55811576, 35696423, 65274791, 56526486	264909, 264769, 264635, 264636
UNCLASSIFIED	reductase	struct	Kinase	UNCLASSIFIED	interleukinrecept
	Contains protein domain (PF00970) - reductase FAD/NAD-binding Cytochrome reductase	Contains protein domain (PF00036) - struct EF hand		Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	Contains protein domain (PF00622) - interleukinrecept
Novel Protein sim. GBank gi 86760 pir A40465 - alanine transaminase (EC 2.6.1.2), cyfosolic - human	Novel Protein sim. GBank gił 1709233 sp P07514 NC5R_BOVIN - NADH- CYTOCHROME B5 REDUCTASE	87800356 (5359, 5360) Novel Protein sim. GBank gil4589604 dbj BAA76824.1 - (AB023197) KIAA0980 protein [Homo sapiens]	90933844 (5361, 5362) Novel Protein sim. GBank gij728837 sp P39194 aLU7_HUMAN - !!!! ALU SUBFAMILY SQ WARNING ENTRY !!!!	Novel Protein sim. GBank gil423468 pir JQ1974 - HTF9-C protein - mouse	87774405 (5365, 5366) Novel Protein sim. GBank gil5114351[gb]AAD40286.1
2678 95001694 (5355, 5356) Novel Protein sim. Gl transaminase (EC 2.0)			90933844 (5361, 5362)		
678	2679	2680	2681	2682	2683

2684	85787151 (5367, 5368) Novel Protein sim. G	Novel Protein sim. GBank gil4886469 emb CAB43385.1 -			264593
2685	88054299 (5369, 5370)	(ALCOUZZA+) Ilybuiteillai piuteili proinio sapietis] 88054299 (5369, 5370) Novel Protein sim. CBank gilj3342729 (AC005331) -		UNCLASSIFIED	
2686		87628690 (5371, 5372) Novel Protein sim. GBank gil4650844 dbj BAA77027.1 - (AB026190) Kelch motif containing protein [Homo sapiens]	Contains protein domain (PF01344) - struct Kelch motif	struct	264259, 29331822, 60432289, 29331827, 29331830, 264909, 264512, 264596, 264769, 264534, 264555, 264556, 264557, 264558, 60170394, 264559, 264486
2687	87998183 (5373, 5374)	87998183 (5373, 5374) Novel Protein sim. GBank gi[5281314[gb AAD41475.1 AF13312 - (AF133123) transcription factor IIIC102 [Homo sapiens]	Contains protein domain (PF00515) - Iranscriptfactor TPR Domain	transcriptfactor	18108394, 18108396, 22278996, 35696286, 22278997, 29331826, 29331828, 66712502, 21906754, 265011, 264760, 264761, 264763, 264689, 21906765, 35696423, 264559, 18108385, 264563
2688	79959584 (5375, 5376)				264908, 264760
2689		94122440 (5377, 5378) Novel Protein sim. GBank gij3880023 emb CAA97339 - Contains protein dom (Z73098) Similarity to yeast hypothetical protein (Swiss Prot Leucine Rich Repeat accession number Q09695); cDNA EST EMBL:D72962 comes from this gene; cDNA EST EMBL:D75947 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00560) - Leucine Rich Repeat		22278997, 22278998, 22278999, 29331824, 35696052, 264906, 264908, 56182435, 264512, 264910, 265009, 60433438, 264512, 264910, 265009, 60433438, 264767, 21906765, 21906766, 21906766, 21906768, 33657182, 27486262, 27486264, 33657349, 18108370, 18108372, 18108374, 35696423, 3569585, 18108385, 22279902
2690		88003055 (5379, 5380) Novel Protein sim. GBank gi[2477513 (AC002398) - F25965_3 [Homo sapiens]	Contains protein domain (PF00620) - struct RhoGAP domain	struct	
2691	91219241 (5381, 5382)	91219241 (5381, 5382) Novel Protein sim. GBank gi 4107276 emb CA467130 - (X98506) acetyl-CoA synthetase (Solanum tuberosum)		synthase	65274572, 56182575, 264259, 29331822, 29331824, 29331826, 29331828, 264112, 265009, 55812038, 264596, 33109954, 265017, 264448, 264388, 264389, 264684
					2000 17, 204110, 204200, 204300, 204001, 21906769, 60170615, 60431528, 55810764, 264634, 264636, 264586, 264637, 22279002, 264564, 264566
2692	94111914 (5383, 5384)	94111914 (5383, 5384) Novel Protein sim. GBank gij3513303 (AC005594) - R26984_1 [Homo sapiens]	Contains protein domain (PF00326) - peptidase Prolyl oligopeptidase family	peptidase	
2693	_			UNCLASSIFIED	264592
2694		94111918 (5387, 5388) Novel Protein sim. GBank gij3122400 sp O35682 MUG_MOUSE - MYELOID UPREGULATED PROTEIN			264559
2692		95345513 (5389, 5390) Novel Protein sim. GBank gil4972740 gblAAD34765.1}- (AF132177) unknown [Drosophila melanogaster]		collagen	35696286, 56994075, 22278999, 264259, 35696052, 29331830, 255011, 264288, 56181562, 264690, 264692, 33657023, 27486262, 263976, 18108376, 35695423, 35695855, 60170394, 83373044, 56526486, 22279000, 22279002, 264566
2696	87874040 (5391, 5392)	87874040 (5391, 5392) Novel Protein sim. GBank gij728831 sp P39188 ALU1_HUMAN - III! ALU SUBFAMILY J WARNING ENTRY II!!		synthase	264594, 21906768, 18108370, 18108372

	-			
35696286, 264259, 29331824, 29331825, 29331826, 29331826, 29331827, 29331828, 35696052, 265006, 265007, 264512, 265009, 60170831, 60433356, 284595, 284758, 81168474, 265010, 265011, 8718859, 264601, 265017, 265011, 8718859, 264601, 265017, 265018, 264761, 18108351, 264448, 264288, 264689, 264689, 21906765, 21906768, 265020, 265021, 60170615, 33557109, 18108376, 35695423, 3569565, 264555, 264558, 60170394, 264559, 18108387, 56526486, 87168518, 22279002, 264553, 264482	22278995, 35696286, 22278996, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331822, 29331824, 29331826, 29331827, 264906, 29331830, 264909, 264906, 29331830, 264909, 264906, 29331830, 265019, 264288, 21906765, 21906765, 21906767, 21906765, 21906767, 21906763, 244150, 33557023, 33657182, 244150, 33657023, 264638, 60432113, 22279000	264768, 18108357, 264690, 264691 60424269, 56182435, 6043229, 60433438, 55811386, 265017, 55811150, 264448, 56181562, 55811957, 264693, 33657109, 60431528, 264629, 55810764, 55811576, 65274791, 60431850, 60432113	18108394, 52645156, 35696286, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 35696052, 29331828, 29146499, 256006, 60433356, 33657084, 18108351, 264288, 264769, 18108359, 21906768, 35695917, 336957023, 27486261, 18108379, 35956423, 18108379, 35956423, 18108382, 83373044, 18108378, 18108388, 18108388, 60432113, 22279000	29331828, 264512, 264555, 264556, 264557, 264558, 264558, 264559, 26332289, 265007, 21906765, 21906768, 265021, 264563
UNCLASSIFIED	Iranscriptfactor	UNCLASSIFIED		UNCLASSIFIED
	Contains protein domain (PF00412) - transcriptfactor			Contains protein domain (PF00642) - Zinc finger C-x8-C-x3-H type (and similar)
2697 91638472 (5393, 5394) Novel Protein sim. GBank gi 5689473 dbj BAA83020.1 - (AB028991) K AA1068 protein [Homo sapiens]	94325891 (5395, 5396) Novel Protein sim. GBank gil841318 (U22818) - mutant sterol regulatory element binding protein-2 (Cricetulus griseus)	87780650 (5397, 5398) 94139836 (5399, 5400) Novel Protein sim. GBank gi[5174395 ref NP_006008.1 pB120 - Brain protein 120	.84148584 (5401, 5402) Novel Protein sim. GBank gil 1572801 (U70854) - F38A5.1 gene product [Caenorhabditis elegans]	57295366 (5403, 5404) Novel Protein sim. GBank gil2605967 (AF030027) - 24 [Equine herpesvirus 4] 87649514 (5405, 5406) Novel Protein sim. GBank gil5689399ldbj BAA82983.1 - (AB028954) KIAA1031 protein [Homo sapiens]
91638472 (5393, 5394)		87780660 (5397, 5398) 94139836 (5399, 5400)		57295366 (5403, 5404) 87649514 (5405, 5406)
2697	2698	2700	2701	2702

2704	2704 A7640516 (6407 6408) Nivial Destrict			20010000 00001000 10001000 001100
<u> </u>	foots 'sous' crosses			264488, 22278995, 22278998, 29331828,
		(AF008554) implantation-associated protein [Rattus		29146499, 264905, 264906, 264907,
		norvegicus]		52644045, 264511, 33657402, 264600,
				264602, 265017, 264605, 264761, 18108351,
				264764, 264687, 264769, 265021, 264691,
				264692, 18108362, 264693, 18108370,
				18108374, 264634, 264635
2705	87771745 (5409, 5410)			264489, 264509, 264511, 264512, 264910,
				264593, 87168474, 264604, 264288, 264687.
				264769, 264638, 264566, 264486
2706	94326789 (5411, 5412) Novel Protein sim.	Novel Protein sim. GBank gij3255952 emb CAA16821.1 -	UNCLASSIFIED	264488, 52646842, 65274572, 22278994,
		(AL021728) /prediction=(method:; /match=(desc:		56994075, 22278997, 264259, 29331824,
		[Drosophila melanogaster]		29331825, 29331826, 29331828, 33656970,
				264907, 264908, 264909, 52644045,
		-		56182435, 265006, 265007, 60433438,
			_	55812038, 21906754, 52644296, 265010.
				264601, 265017, 265019, 264681, 264448,
				264682, 264288, 264686, 264687, 264688,
			•	21906766, 21906769, 55811957, 35695917.
				265020 265021 60170615 264690 264691
				33657023 264692 264693 65274620
				22486264 26362; E01080; CCE 10EC;
				2.100201; 2003; 2, 101003; 1, 1010003; 1, 1010003; 1, 101003; 1, 101003; 1, 101003; 1, 101003; 1, 101003; 1, 1
				83373044 65274727 87168518 22279000
2707	88089839 (5413, 5414) Novel Protein sim.	Novel Protein sim. GBank gij3417294 (AC004381) -		22278996, 22278998, 56182435, 21906754
				87168559, 265017, 264448, 52645129
2708	91011351 (5415, 5416) Novel Protein sim.	Novel Protein sim. GBank gi 545790 bbs 147178 - DARPP-	UNCLASSIFIED	65274572, 264259, 29331822, 29331825,
		32=dopamine and cAMP-regulated phosphoprotein (human,		60432289, 29331826, 29331827, 29331828,
		brain, Peptide, 204 aa]		264909, 264510, 265007, 264910, 60433356,
				60433438, 33109954, 265010, 265011,
	_			264369, 264288, 264765, 264693, 264565
2709	94853988 (5417, 5418) Novel Protein sim.	Novel Protein sim. GBank gij3169705 (AC004780) -	UNCLASSIFIED	29331822, 18108370, 18108374, 83373044
	_	F17127_1 [Homo sapiens]		
2710	87627979 (5419, 5420) Novel Protein sim.	Novel Protein sim. GBank gij4468311 emb CAB37992 -		29331824, 264759, 264693, 18108382,
		(AL031432) dJ465N24.2.1 (PUTATIVE novel protein)		18108388
		(isoform 1) [Homo sapiens]		

FIED	760515) - Iransferase 264488, 35696286, 22278998, 264259, 29331824, 60432289, 35696052, 264508, 224906, 66712502, 5264045, 265006, 6043229, 33657402, 60433229, 33657402, 6043322, 265016, 21906766, 21906766, 21906766, 21906766, 21906769, 55811957, 35699517, 265020, 265021, 60170615, 264691, 264692, 33657023, 33657109, 33657109, 33657182, 33657349, 18108370, 18108374, 35696423, 3569855, 264555, 52644332, 65182323, 87168518, 60432113	UNCLASSIFIED 66714117, 264906, 264563		glycoprotein 264091, 264259, 29331822, 66/14117, 264369, 264363, 264563, 264563, 264558	264693 UNCLASSIFIED 29331822, 87168559, 265019, 265021, 5264150, 264691
N MOUSE - MYELOID	Novel Protein sim. GBank gi 5081315 gb AAD39343.1 AF07660 - (AF076607) TPR Domain prediabetic NOD sera-reactive autoantigen [Mus musculus]	88003064 (5425, 5426) Novel Protein sim. GBank gil2477513 (AC002398) - F25965_3 [Homo sapiens] 13228218 (5427, 5428)	94122454 (5429, 5430) Novel Protein sim. GBank gil4321968 gb AAD15897 - (AF067430) Smarce1-related protein [Mus musculus]	ocuusuoo (5431, 5434) Novel Protein sim. Gbank gilz47/513 (ACU02398) - F25965_3 [Homo sapiens] 80077461 (5433, 5434) Novel Protein sim. GBank gil3327046[db][BAA31591] - (AB014516) KIAA0616 protein [Homo sapiens]	79604062 (5435, 5436) 88180423 (5437, 5438) Novel Protein sim. GBank gil746495 (U23515) - weakly similar to gastrula zinc finger protein [Caenorhabditis
94111920 (5421, 5422)	94312071 (5423, 5424)	2713 88003064 (5425, 5426) Novel 2714 13528218 (5427, 5428)		2710 000U3U00 (2431, 5434) NOVER 2717 80077461 (5433, 5434) Novel (ABD)	2718 79604062 (5435, 5435) 2719 88180423 (5437, 5438) Novel Pro similar to

2720	95086242 (5439, 5440)	2720 95086242 (5439, 5440) Novel Protein sim. GBank gil 1335873 (U46690) - ATP-dependent RNA helicase [Mus musculus]	Contains protein domain (PF00270) - Ihelicase DEAD/DEAH box helicase	helicase	18108374, 60424179, 264489, 56182435, 21906765, 21906766, 35696423, 22278997, 265020, 265022, 265006, 265008, 264092, 264636, 60432229, 264691, 264692, 33657023, 264693, 33657402, 83373044, 29331824, 18108386, 5264519, 21906754, 3569052, 29331828, 87168474, 264100, 265010, 265011, 265019, 22279002, 264905, 264482, 26453, 264906, 18108351, 264681, 55644145, 264904, 18108370, 264908, 66712502, 55644045, 264904, 264904, 18108370, 264904, 26490
2721	95345523 (5441, 5442) Novel Protein sim. gil4929663 gb AAC protein [Homo sapi	Novel Protein sim. GBank gil4929663 gb AAD34092.1 AF15185 - (AF151855) CGI-97 protein [Homo sapiens]	Contains protein domain (PF01172) - Uncharacterized protein family UPF0023		22278995, 35696286, 264259, 29331822, 29331824, 66714117, 29331826, 264906, 60433438, 265017, 18108351, 264448, 264288, 264769, 21906766, 265021, 33657109, 263969, 60431528, 284629, 55811576, 65274791, 35695855, 264631, 264637, 60170394, 56182323, 22279000
2722	91638807 (5443, 5444) Novel Protein sim. (AC004997) match AA165465 (NID:g1 T84026 (NID:g712: including: AF04065 (PID:g2104571), A	Novel Protein sim. GBank gij3212997]gbJAAC23434.1] - (AC004997) match to ESTs AA667999 (NID.g2626700), AA165465 (NID.g1741481), Z45871 (NID.g575105), and T84026 (NID.g712314); similar to various tre-like proteins including: AF040654 (PID.g2746883), D13644 (PID.g2104571), AL02114	Contains protein domain (PF00566) - oncogene TBC domain	опсоделе	35696286, 22278999, 21906754, 265017, 264762, 264288, 21906765, 21906767, 21906768, 35695917, 18108362, 27486262, 35695855, 264558, 264559
2723	87387732 (5445, 5446)			UNCLASSIFIED	264508, 264509, 264906, 264909, 264910, 55812038, 264768, 264687, 264629, 264636, 264486
2724	87639563 (5447, 5448) Novel Protein sim. gi 4680681 gb AAD protein [Ното sapi	Novel Protein sim. GBank gil4680681lgblAAD27730.1 AF13295 - (AF132955) CGI-21 protein [Homo sapiens]		ubiquitin	18108396, 22278999, 20281099, 29331824, 29331826, 60432289, 29331828, 60170831, 60432229, 60433438, 18108351, 264682, 21906766, 21906767, 21906769, 35695917, 33657023, 33657029, 18108372, 18108374, 35695855, 22279000, 22279002
2725	2725 94853991 (5449, 5450) Novel Protein sim. F17127_1 [Homo s	Novel Protein sim. GBank gij3169705 (AC004780) - F17127_1 [Homo sapiens]		UNCLASSIFIED	264488, 52644507, 264259, 29331827, 21906754, 265011, 18108351, 26448, 264288, 264685, 264689, 35695917, 265020, 33657182, 27486261, 18108370, 18108374, 35696423, 18108385, 22279000
2726	86880599 (5451, 5452) Novel Protein sim. R26660_1, partial	l Novel Protein sim. GBank gij3342738 (AC005328) - R26660_1, partial CDS [Homo sapiens]		мнс	264488, 264828, 264685

	ATOTO (2453, 5454) Novel Prolein sim. Gi gij731267 sp 23219 SUBUNIT PSEUDOL (PSEUDOURIDYLAT HYDROLYASE)	Novel Protein stim. GBank gij731267jsplP39219fILUA_ECOLI - RIBOSOMAL LARGE SUBUNIT PSEUDOURIDINE SYNTHASE A (PSEUDOURIDYLATE SYNTHASE) (URACIL HYDROL YASE)	Contains protein domain (PF00849) - synthase RNA pseudouridylate synthase	6274572, 56182575, 22278994, 56994075, 22278996, 22278996, 22278999, 60432049, 29331822, 26331822, 26331828, 36574417, 29331826, 29331827, 35696052, 29331828, 3365970, 264509, 66712502, 264910, 33657402, 60433438, 264756, 264910, 3365709, 264765, 21906765, 21906767, 21906767, 21906765, 21906767, 21906765, 21906767, 21906765,
2728		Novel Protein sim. GBank gij3880433[emb]CAA91399] - (286521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL:C09217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - UNCLASSIFIED Mitochondrial carrier proteins	264557, 2931822, 264559, 26448, 264288
2729	94126024 (5457, 5458)	94126024 (5457, 5458) Novel Protein sim. GBank gij2408095 emb CAB16300 - (299168) putative RNA splicing protein [Schizosaccharomyces pombe]	Contains protein domain (PF00153) - UNCLASSIFIED Mitochondrial carrier proteins	65274572, 264259, 60432289, 66712502, 56182435, 264448, 264288, 264369, 55811957, 265021, 264557, 60432113
2730	94126026 (5459, 5460)	94126026 (5459, 5460) Novel Protein sim. GBank gij3860433jemb CAA91399 - (Z66521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL:C09217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - transport Mitochondrial carrier proteins	264887, 29331824, 29331826, 35696052, 264107, 56182435, 265008, 265009, 264592, 60431735, 265001, 264601, 266017, 18108351, 264288, 29148627, 55811957, 264557, 264590, 18108368, 18108374, 264557, 264558, 264559, 18108387, 56526486, 264566, 264486
2731		87723022 (5461, 5462) Novel Protein sim. GBank gij1723239jspj010166jYAUB_SCHPO - HYPOTHETICAL 35.7 KD PROTEIN C26A3.11 IN CHROMOSOME I	Contains protein domain (PF00795) - Carbon-nitrogen hydrolase	264259, 35696052, 265006, 264758, 264762, 264448, 264288, 29148627, 21906769, 87168518, 22279002
2732		94126028 (5463, 5464) Novel Protein sim. GBank gij3860433jemb CAA91399 - (266521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL.C09217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - UNCLASSIFIED Mitochondrial carrier proteins	264687, 264489, 18108358, 56182435, 264689, 35696423, 55811957, 265021, 265006, 265008, 264910, 265009, 264690, 264555, 264259, 264556, 264557, 264558, 264559, 18108383, 33657109, 87168518, 265010, 264601, 60432113, 265017, 264905, 264448, 263972, 264369, 264567
2734	87363060 (5467, 5466) 94140286 (5467, 5468) Novel Protein sim. Gl (AB017614) OASIS p	Novel Protein sim. GBank gil4519621 dbj BAA75670.1 - (AB017614) OASIS protein [Mus musculus]	UNCLASSIFIE Contains protein domain (PF00170) - transcriptfactor bZIP transcription factor	29331825, 264509, 264909 60424179, 52644507, 56182575, 264259, 29331828, 264907, 264510, 264910, 60433356, 265019, 55811150, 264681, 264763, 264687, 33657182, 18108370, 60431528, 60431850, 56182323, 83373044

2735	87712336 (5469 5470)	2735 (87712336 (5469 5470) Novel Protein eim GBank wijgschege (ACOOE278) = ET.	3.		
<u> </u>	,	gb/T21276, gb/T45403, and gb/A4586113 come from this		giycopratein	222/8996,60432289,29331827,29146498, 264108 264000 264112 32657402
		gene. [Arabidopsis thaliana]			87168474 265017 264762 26444B 264764
		•			264684 21006765 264603 23657100
					263976 264636 264638 264557 22270000
					22270002 264567
2736	80247655 (5471, 5472)			INCI ASSIBIED	26400E 264628 264620 262020 264622
					264564
2737	87604526 (5473, 5474)	87604526 (5473, 5474)			264690
2738	85731808 (5475, 5476)	Novel Protein sim. GBank gi[2558501 dbj]BAA22896 -			264488, 265009, 264768, 264691
		(D63850) hepatoma-derived growth factor [Mus musculus]			
2739	94319834 (5477, 5478)	94319834 (5477, 5478) Novel Protein sim. GBank gi[5420387 emb CAB46679.1 -		UNCLASSIFIED	264684 83373044 264566
		(AJ243459) proteophosphoglycan [Leishmania major]			
2740		Novel Protein sim. GBank gi 3417386 emb CAA75495 -		UNCLASSIFIED	264488, 56182575, 22278995, 35696286
		(Y15197) microtubule-associated protein, MAP-115 [Mus			22278997, 22278998, 22278999, 264259
		[musculus]			29331822, 29331824, 29331825, 29331827
					35696052, 29331828, 29146498, 29331830
					265006, 265007, 265009, 60432229,
_					33657402, 55812038, 87168474, 265010.
					265011, 265017, 265018, 265019, 264605.
					264681, 264288, 264369, 52644229.
					21906765, 21906766, 21906767, 21906768
					21906769, 265020, 265022, 264691, 264692
	_				33657109, 18108370, 18108374, 55810764
					35695855, 264634, 60431850, 264639
					56182323 18108382 18108385 65274727
					22279002, 264564
2741	88047518 (5481, 5482)	88047518 (5481, 5482) Novel Protein sim. GBank gij3242764 (AC005154) - similar		UNCLASSIFIED	22278996 52644045 52644229 2190676R
		to protein U28928 (PID:9861306) [Homo sapiens]			21906769, 265020, 60170615, 264691
2742		87648644 (5483, 5484) Novel Protein sim. GBank	Contains protein domain (PF00652) - transferase		264259, 264905, 264758, 55812038, 264369
		gil4758412[ref[NP 004472.1]pGALN - UDP-N-acetyl-alpha-			29148627
		D-galactosamine:polypeptide N-			
		acetylgalactosaminyltransferase 2 (GalNAc-T2)			
2743	87627991 (5485, 5486)	Novel Protein sim. GBank gil4468311jemb CAB37992 -		UNCLASSIFIED	35696286, 264259, 264906, 264908, 265006.
		(AL031432) dJ465N24.2.1 (PUTATIVE novel protein)			60433438, 265017, 18108351, 264448.
		(isoform 1) [Homo sapiens]			264764, 264288, 21906765, 21906767,
					264690, 264691, 264693, 263969, 263971
					35695855, 264637, 264558, 18108382
					60432113
2744	94126030 (5487, 5488)	94126030 (5487, 5488) Novel Protein sim. GBank gij3880433jemb CAA91399-	Contains protein domain (PF00153) - kinase		18108374, 264488, 56182435, 21906765,
		(Z66521) similar to mitochondrial RNA splicing MSR4 like	Mitochondrial carrier proteins		35696423, 35695917, 35695855, 265020,
		protein; cDNA EST EMBL:C09217 comes from this gene			265021, 264511, 265009, 264490, 264556,
		[Caenorhabditis elegans]			264259, 264557, 56182323, 264558, 264559,
					18108383, 29331824, 18108385, 33657109,
_					29331826, 21906754, 29331827, 29331828,
					33657349, 87168518, 265018, 264905,
					264482, 264448, 264486, 264369, 264288

2745	87740125 (5489, 5490)	2745 87740125 (5489, 5490) Novel Protein sim. GBank gil405795 gb AAD19826 - (AF038963) RNA helicase [Homo saplens]	Contains prolein domain (PF00271) - helicase Helicases conserved C-terminal domain	helicase	36996286, 264509, 264905, 264907, 264908, 264909, 264909, 264510, 264512, 265008, 264758, 264601, 265017, 264604, 264763, 264288, 264686, 264769, 264693, 35696423, 35695855, 264634, 264636, 264563, 264564, 264656
2746	95418601 (5491, 5492)	95418601 (5491, 5492) Novel Protein sim. GBank gil4758738 ref NP_004680.1 pMTA1 - metastasis associated 1	Contains protein domain (PF00320) - UNCLASSIFIED GATA zinc finger	UNCLASSIFIED	22278996, 22278998, 22278999, 29331822, 29331826, 28331827, 35696052, 29331828, 264905, 264905, 264908, 264908, 264909, 264909, 264909, 265812038, 265018, 264683, 264288, 21906765, 21906765, 21906769, 264690, 33657109, 18108368, 18108374, 264558, 18108385, 22279000, 264563
2747	94112677 (5483, 5494)			glycoprotein	264569, 52644507, 18108394, 22278995, 35696286, 22278997, 22278999, 52645080, 29331824, 26545080, 29331824, 26545080, 25696052, 284907, 224908, 264909, 265009, 33696052, 284907, 265409, 265009, 37108547, 265010, 264768, 21906769, 35695917, 60170615, 264692, 33657023, 52645129, 27488264, 60431528, 18108374, 35696423, 35695855, 264565, 56182323, 18108385, 264482
2748	91214983 (5495, 5496)	2748 91214983 (5495, 5496) Novel Protein sim. GBank gi[4191272 emb CAA09984 - (AJ012295) apaG protein [Rhizobium etli]	Contains protein domain (PF00646) - F-box domain.		65274572, 29331828, 264112, 264511, 265019, 264760, 264767, 264768, 264769, 21906768, 21906769, 265020, 27486262, 56526486, 87168518, 22279000
2749	87346307 (5497, 5498)				264259, 264908, 264510, 265008, 265009, 264760, 264369, 264768, 264563
2750	87336344 (5499, 5500)	87336344 (5499, 5500) Novel Protein sim. GBank gil1872498 (U74297) - PiUS [Oryctolagus cuniculus]		UNCLASSIFIED	264488, 52644507, 18108396, 56994075, 264259, 29331825, 29331826, 29331827, 29331828, 264508, 2655009, 264910, 264591, 264595, 33657084, 265011, 265019, 18108351, 264288, 264686, 264769, 264689, 55811957, 264639, 27486264, 18108370, 18108374, 264558, 18108385, 264482, 264563
2751		87057465 (5501, 5502) 88062675 (5503, 5504) Novel Protein sim. GBank gij3041859 (AC004534) - OG-2 homeodomain protein-like; similar to U65067 (PID:g1575526) [Homo sapiens]	UNCLASS Contains protein domain (PF00046) - homeobox Homeobox domain	UNCLASSIFIED homeobox	29331822, 29331824, 265017, 33657023

						-	r			
56182575, 56994075, 22278998, 29331822, 29331824, 29331825, 29331826, 265007, 264593, 55812038, 33109954, 18108351, 264288, 56181562, 21906767, 21906768, 265021, 264693, 18108374, 65274791, 264632, 56182323, 22279002, 264563, 264567	22278999, 66714117, 29331827, 3569052, 29331828, 284906, 264908, 264909, 265011, 265017, 265018, 265019, 264288, 21906765, 21906767, 265022, 33657023, 264693, 26182323, 18108382, 22279000	, , , , , , , , , , , , , , , , , , , ,	264034 85658542, 264693	264693	65274572, 35696286, 66714117, 29331828, 264508, 56182435, 21906754, 55811957, 264629, 264636, 56182323, 22279002	264908	22278998, 29331822, 29331830, 265010, 265019, 265019, 264288, 21906765, 21906768, 21906769, 265020, 56182323, 22279002, 264563	18108394, 22278998, 264906, 264909, 265006, 265007, 264757, 265010, 265011, 265017, 265019, 18108351, 264481, 264689, 264686, 264691, 18108362, 264693, 18108368, 18108370, 18108381, 18108362, 18108382, 18108381, 18108382, 1810	264369, 35696423	52646842, 264259, 29331822, 29331825, 29331826, 29331826, 29331828, 33656970, 264905, 264907, 29331830, 265006, 265009, 21906754, 265019, 264448, 21906769, 27486262, 56182323, 56526486, 87168518, 264487
UNCLASSIFIED	struct	transcriptfactor	eph eph	struct		UNCLASSIFIED	nuclease	. struct	- transcriptfactor	
Contains protein domain (PF00646) - F-box domain.	Contains protein domain (PF00036) - EF hand	Contains protein domain (PF00096) - transcriptfactor Zinc finger, C2H2 type	UNN Contains protein domain (PF00017) - eph Src homology domain 2				Contains protein domain (PF00023) - nuclease Ank repeat	Contains protein domain (PF00989) - struct PAS domain	Contains protein domain (PF01352) - transcriptfactor KRAB box	Contains protein domain (PF00370) - FGGY family of carbohydrate kinases
Novel Protein sim. GBank gij3851648 (AF098301) - neural F Contains protein domain (PF00646) - UNCLASSIFIED box protein NFB42 [Rattus norvegicus]	94115513 (5507, 5508) Novel Protein sim. GBank gi[535428 (U13736) - calmodulin- Contains protein domain (PF00036) - struct [Ike protein [Pisum sativum] EF hand	88001472 (5509, 5510) Novel Protein sim. GBank gi[2996653 (AC004510) - R30385_2 [Homo sapiens]	11465908 (5511, 5512) 95361590 (5513, 5514) Novel Protein sim. GBank gil 1173539 (U30473) - putative g5361590 (5513, 5514) Novel Protein sim. GBank gil 1173539 (U30473) - putative src-like adapter protein; non-catalytic src-like adapter protein containing SH3 and SH2 domains; homolog of mouse SLAP; Method: conceptual translation supplied by author Homos sapiens!	79637846 (5515, 5516) Novel Protein sim. GBank gi]3608372 (AF053768) - brain specific cortactin-binding protein CBP90 [Rattus non-egicus]	91005312 (5517, 5518) Novel Protein sim. GBank gi 2072200 (U94863) - p40 [Boma disease virus]		Novel Protein sim. GBank gil4914573 emb CAB43685.1 - (AL050390) hypothetical protein [Homo sapiens]	Novel Protein sim. GBank gij3136150 (AF050182) - PERIOD 3 [Mus musculus]	87539968 (5525, 5526) Novel Protein sim. GBank gi[3511122 (AF060503) - zinc	94305140 (5527, 5528) Novel Protein sim. GBank gi[2905643 (AF045244) - ribitol kinase [Klebsiella pneumoniae]
2753 94138972 (5505, 5506) Novel Protein sim. Gi	94115513 (5507, 5508) [88001472 (5509, 5510) I	11465908 (5511, 5512) 95361590 (5513, 5514) N 8	79637846 (5515, 5516)	91005312 (5517, 5518)	79824798 (5519, 5520)	87639597 (5521, 5522) Novel Protein sim. G (AL050390) hypothel	87592699 (5523, 5524) Novel Protein sim. G PERIOD 3 [Mus mus	87539968 (5525, 5526)	94305140 (5527, 5528)
2753 (2754		2757	2758	2759		2761	2762	2763	2764

	8, 5, 6, 4			ດັວບໍ່ 4, ດຸ
264488, 65274572, 22278995, 35696286, 22278998, 264259, 29331822, 29331824, 35696052, 264259, 29331822, 29331824, 35696052, 264269, 264509, 264905, 264909, 264907, 264908, 264909, 264512, 264910, 265009, 264592, 264595, 264763, 264763, 264763, 264763, 264764, 264764, 264763, 264764, 264764, 264764, 264696, 264764, 264764, 264696, 264764, 264696, 265018, 26501, 264604, 21906764, 21906764, 21906764, 21906764, 21906764, 264691, 264691, 264692, 36696423, 35695855, 264631, 264634, 264636, 26463	264488, 65274572, 60432289, 264907, 264488, 65274572, 60432289, 264907, 264685, 264689, 35695917, 265022, 264693, 264628, 65274791, 264635, 264555, 264556, 264527, 264638, 264558, 264559, 83373044,	264112, 263974, 264558 264488, 263994, 264489, 35696286, 22278998, 22278999, 264259, 29331822,	60432289, 29331826, 35696052, 264508, 264509, 264905, 264906, 264907, 29331830, 264908, 264909, 264510, 264511, 264512, 265008, 264910, 265009, 264591, 264592, 264593, 33657402, 264594, 264595, 264757, 264596, 265011, 264604, 265019, 264760,	264681, 18108351, 264764, 264288, 264766, 264686, 264766, 264686, 21906766, 21906766, 21906769, 3565931, 265021, 265022, 33657023, 33657109, 265628, 264629, 18108374, 35696423, 35693655, 264630, 264631, 264636, 264636, 264536, 264564, 264565, 264565, 264565, 264566, 264565, 264566, 264565, 264566, 264566, 264566, 264567, 264688, 264567
	dna_ma_bind	struct UNCLASSIFIED		
	Contains protein domain (PF00076) - dna_ma_bind RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)			
2765 94315105 (5529, 5530) Novel Protein sim. GBank gij4688672 emb GAA17688.2 - (AL022018) /prediction=(method:: /match=(desc: [Drosophila melanogaster]	Novel Protein sim. GBank gijs441611jembjCAB46854.1j (AJ388555) hypothetical protein [Canis familiaris]	94322238 (5535, 5536) Novel Protein sim. GBank gil 54322[emb CAB46721.1] - (AL031427) dJ167A19.1 (novel protein) [Homo sapiens]		
·65 94315105 (5529, 5530)		2768 94322238 (5535, 5536) N		
51	57	27		

	т		T		
264488, 56182575, 264259, 29331822, 29331824, 66714117, 29331825, 60432289, 29331829, 2645908, 29331830, 56182435, 264592, 33657402, 264448, 264369, 264288, 60170615, 264691, 33657023, 264692, 33657109, 18108374, 55811576, 264634, 264636, 56182323, 83373044, 60432113	18108398, 22278995, 22278999, 22278999, 224105, 265006, 265019, 18108351, 264687, 21906765, 18108364, 264629, 18108374, 264631, 18108388, 18108388	264259, 29331822, 29331824, 29331825, 264369	264259, 29331822, 29331824, 66714117, 60432289, 29331827, 264905, 265009, 264592, 55812038, 65274444, 264766, 21906769, 33657109, 263976, 264555, 264638, 264557, 83373044, 264563, 264564	18108374, 264686, 264687, 263976, 56182435, 264689, 55810764, 21906766, 55186423, 264689, 55810764, 21906766, 55186423, 55811576, 65274791, 56181686, 55811957, 35695855, 264110, 265021, 265008, 264092, 264094, 60431850, 264637, 264691, 264697, 264697, 264697, 264697, 264697, 264177, 264181, 55812038, 29331825, 60424269, 18108385, 26331826, 29331826, 28331826, 28331826, 28331826, 284107, 60432113, 265017, 5581136, 264884, 264681, 264906, 18108370, 264884, 264683, 264764, 264289, 264684, 264766, 263974	22278995, 35666286, 22278996, 22278996, 226425, 29331826, 60432289, 35696052, 264112, 33657402, 21906754, 87168559, 265017, 265018, 224289, 21906767, 21906769, 35695917, 265020, 265021, 33657023, 33657109, 18108370, 263976, 35696423, 35695855, 87168518, 222799000, 264482
tubulin	kinase	UNCLASSIFIED	histone	UNCLASSIFIED	ribosomalprot
	Contains protein domain (PF00400) - kinase WD domain, G-beta repeat		Contains protein domain (PF00850) - Histone deacetylase family		Contains protein domain (PF00380) - ribosomalprot Ribosomal protein S9/S16
2769 95311088 (5537, 5538) Novel Protein sim. GBank gil5419859 emb CAB46375.1 - (AL096725) hypothetical protein [Homo sapiens]	Novel Protein sim. GBank gi 5701965 emb CAB52157.1 - (AL109736) WD repeat protein (Schizosaccharomyces pombe]	88084071 (5541, 5542) Novel Protein sim. GBank gij3093433 (AC004125) - Unknown gene product [Homo sapiens]	95357309 (5543, 5544) Novel Protein sim. GBank gij4885531 ref NP_005465.1 pNY Contains protein domain (PF00850) - histone C- histone deacetylase 5 Histone deacetylase family	94138994 (5545, 5546) Novel Protein sim. GBank gij3288888 (AC005253) - R26445_1 [Homo sapiens]	87819906 (5547, 5548) Novel Protein sim. GBank gil465852lsplp34388 YLS3_CAEEL - HYPOTHETICAL 70.7 Ribosomal protein S9/S16 KD PROTEIN F09G8.3 IN CHROMOSOME III
95311088 (5537, 5538)	2770 87730182 (5539, 5540) Novel Protein sim. GB (AL109736) WD repea pombe]	88084071 (5541, 5542)			
2769	2770	2771	2772	27773	2774

264488, 22278995, 56994075, 22278996, 35695286, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331827, 29331824, 29331825, 265007, 21906754, 265017, 265019, 264448, 264682, 264369, 21906765, 21906766, 21906767, 21906768, 21906769, 3699517, 265021, 265022, 60170615, 264991, 18108370, 35696423, 65274791, 3569585, 264563, 264486, 264486, 264567, 264486, 264486, 264486, 264567, 264486, 264486, 264486, 264567, 264486, 264486, 264567, 264486	56182575, 22278998, 22278999, 264259. 29331822, 29331824, 264908, 29331830, 264510, 33657402, 21906754, 55811386. 265017, 265019, 26448, 264288, 21906765, 21906766, 21906767, 21906768, 21906768, 21906768, 21906768, 258020, 265020, 265021, 265022, 60170615, 53810764, 55811576, 264555, 56526486, 22220, 20170615, 264555, 264555, 26526486, 22220, 20170615, 264555, 264	П		UNCLASSIFIED 22278997, 264259, 29331824, 66714117, 35696052, 265006, 264512, 264448, 264288, 29148627, 18108364, 20281149, 18108370, 264629		iltin 60424179, 29331824, 60424269, 66714117, 29331826, 56182435, 87168474, 265017, 264764, 56181562, 21906765, 21906766, 21906768, 35695917, 265020, 33657023, 35695855, 56182323, 87166518	
Contains protein domain (PF00177) - ribosomalprot Ribosomal protein S7p/S5e		OND	ONO		Contains protein domain (PF00780) - kinase CNH domain	Contains protein domain (PF00443) - ubiquitin Ubiquitin carboxyl-terminal hydrolase family 2	Contains protein domain (PF00560) - struct Leudine Rich Repeat
Novel Protein sim. GBank gi 4689132 gb AAD27775.1 AF07704 - (AF077042) 30S ribosomal protein S7 homolog [Homo sapiens]				ord+97.29 (35357, 3538) Novel Protein sim. GBank gi[4680711[gb]AAD27745.1[AF13297 - (AF132970) CGI-36 protein [Homo sapiens]	94679397 (5559, 5560) Novel Protein sim. GBank gil4758524 ref NP_004825.1 pHGK - HPK/GCK-like kinase	Novel Protein sim. GBank gil4469352[gb AAD21222] - (AF069502) ubiquitin specific protease UBP43 [Mus musculus]	94233146 (5563, 5564) Novel Protein sim. GBank gil4505013[ref NP_002310.1 pLRN - leucine-rich neuronal protein
	2776 87791557 (5551, 5552)	2777 79818729 (5553, 5554) 2778 82112411 (6666 6668)	2770 87640700 (EEET EEED)	0,049/29 (555/, 5558)			2782 94233146 (5563, 5564)

2783	80016629 (5565, 5566)	Novel Protein sim. GBank gif728831 sp P39188 ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII		tm7	264909, 264628, 263978, 263981
2784	$\overline{}$				264259, 29331822, 29331824, 29331825, 264482
2785		88071930 (5569, 5570) Novel Protein sim. GBank gi 2134933 pir S58890 - collapsin response mediator protein - human		UNCLASSIFIED	22278999, 264908, 264758, 265018, 264769, 21906765, 21906768, 21906769, 265020, 264564
2786		Bank gi 2073564 (U80223) - eukaryotic alpha kinase; DGCN2 [Drosophila	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	264905, 265017
2787	87901266 (5573, 5574)	Novel Protein sim. GBank gi 5174507 ref NP_006020.1 pMA1 - paraneoplastic neuronal antigen		UNCLASSIFIED	264768, 21906765, 21906768, 55811957, 22278999, 264093, 60170615, 264259, 29331822, 33687109, 29331827, 35696052, 264100, 264105, 264908, 263977, 55811576, 264635, 264637, 26279002, 55811150, 264369, 264288
2788		Novel Protein sim. GBank gi[3252826 (AC004382) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	
2789	85491275 (5577, 5578)	85491275 (5577, 5578) Novel Protein sim. GBank gi[2495729 sp Q92556 Y281_HUMAN - HYPOTHETICAL PROTEIN KIAA0281 (HA6725)		UNCLASSIFIED	264103, 21906769, 264693
2790	87602784 (5579, 5580)	Novel Protein sim. GBank gi 5101772 emb CAB45135.1 - (AJ242978) p621 [Homo sapiens]		struct	264488, 264769, 264910, 264631, 254636, 264690, 264691, 264259, 264638, 29331824, 264508, 264509, 264905, 264563, 264906, 264628, 18108370, 264907, 264764, 264908, 264288, 264909
2791	j	88083195 (5581, 5582) Novel Protein sim. GBank gil2911266 (AC002550) - Unknown gene product [Homo sapiens]			21906764, 18108368
2792					22278996, 22278997, 264259, 29331822, 29331824, 29331824, 29331825, 66714117, 29331826, 60432289, 29331827, 35696052, 264906, 6643229, 66433356, 6043348, 33109954, 265010, 265011, 265017, 265018, 265019, 26448, 264289, 26448, 265022, 265021, 33657023, 263974, 18108374, 65274791, 3565523, 284556, 264558, 264568, 264558, 264558, 264558, 264538, 264368, 264558, 264558, 264558, 264523, 25279000, 264567
2793				UNCLASSIFIED	264259, 60432289, 66712502, 265009, 264636
2794	85794830 (5587, 5588)			UNCLASSIFIED	264689, 265022, 264691, 18108388, 264567

6 4 4 4 6 4 6 4 4 5 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7		
UNCLASSIFIED	UNCLASSIFIE	UNCLASSIFIED
Contains protein domain (PF00179) - ubiquitin Ubiquitin-conjugating enzyme UNCLA	Contains protein domain (PF00569) - UNCLASSIFIED Zinc finger present in dystrophin, CBP/p300	
Bank T715.1 AF 13294 - (AF 132940) CGI-06 ns]	95110790 (5593, 5594) Novel Protein sim. GBank gil4838557[gb]AAD31040.1] - (AF143859) potassium channel modulatory factor DEBT-91 [Mus musculus]	86198005 (5595, 5596) Novel Protein sim. GBank gi 2852645 (AF007160) - unknown [Homo sapiens]
95334888 (5589, 5690)		86198005 (5595, 5596)
2795	2797	2798

22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331826, 35696052, 29331828, 264909, 6043336, 33657402, 33109954, 87168474, 264448, 52644229, 21906766, 21906767, 21906768, 3569517, 265020, 265021, 265022, 52644150, 35695855, 264634, 60432113, 22279000	264488, 56994075, 264259, 20281099, 29331825, 29331827, 264905, 56182455, 265006, 265011, 87168599, 265017, 265019, 264488, 264288, 264766, 264686, 60170615, 264631, 264692, 27486255, 264628, 264629, 264636, 264567, 264558, 264559, 87168518, 264564, 264566, 264567	265007, 264687	264448, 35695855	264639	264566	264092, 264259, 29331822, 29331824, 29331827, 29331828, 264508, 265007, 265009, 21906754, 264369, 264288, 264632, 60170394, 264563, 264482	22278995, 22278996, 22278999, 264259, 29331822, 29331824, 29331826, 29331827, 29146498, 265008, 265009, 60433438, 265017, 265018, 265019, 264448, 264288, 21906765, 21906767, 21906769, 21906769, 29148629, 265022, 52644150, 56182323
UNCLASSIFIED	glycoprotein		transport	UNCLASSIFIED	peptidase		UNCLASSIFIED
		Contains protein domain (PF00627) - UBA domain	Contains protein domain (PF00083) - Sugar (and other) transporter		Contains protein domain (PF01585) - peptidase G-patch domain		
2799 88090651 (5597, 5598) Novel Protein sim. GBank gi 3252825 (AC004382) - Unknown gene product [Homo sapiens]	88316481 (5599, 5600) Novel Protein sim. GBank gi(4240301 dbj BAA74929.1 - (AB020713) KIAA0906 protein [Homo saplens]	5602)	88082477 (5603, 5604) Novel Protein sim. GBank gi 2337865 (AC002464) - organic Contains protein domain (PF00083) - Iransport cation transporter; 50% similarity to JC4884 (PID:g2143892) Sugar (and other) transporter [Homo sapiens]	(9095	2804 57111131 (5607, 5608) Novel Protein sim. GBank gil4559368 gb AAD23029.1 AC00658 - (AC006585) hypothetical protein [Arabidopsis thaliana]	5610)	87898951 (5611, 5612) Novel Protein sim. GBank , gi 1168973 sp P44403 CLPB_HAEIN - CLPB PROTEIN
9 88090651 (5597, 55		2801 86068814 (5601, 5602)	2802 88082477 (5603, 56	3 79577446 (5605, 56	4 57111131 (5607, 56	2805 87398486 (5609, 5610)	2806 87899951 (5611, 56

52644507, 52645156, 52646842, 18108398, 56182575, 22278994, 22278995, 22278996, 56994075, 35696286, 22278997, 22278996, 22278999, 22278999, 22278999, 22278999, 22278999, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 35696052, 29331828, 33656970, 264100, 264105, 264907, 25640405, 6043336, 264594, 6043348, 52646917, 21906764, 21906764, 265011, 8716859, 265017, 265018, 265019, 18108351, 264448, 264763, 26487, 265021, 265022, 52644150, 33657023, 52645129, 33657109, 33657102, 27486262, 27486265, 27486265, 27486265, 27486265, 27486265, 27486265, 27486265, 35695763, 18108377, 35696423, 35695855, 52644332, 83373044, 18108377, 35696423, 35695855, 52644332, 83373044, 18108337, 35695855, 52644332, 83373044, 18108385, 18108387, 87168518, 60437113	60432289, 264510, 265010, 265017, 265018, 264681, 264686, 265021, 264690, 22279000, 264566	ed 18108351	264569, 22278996, 264091, 264259, 29331828, 29146499, 29148629, 29148784, 264693	22278999 264259 66712502 264693	264106	56182575, 29331828, 264906, 66712502, 55811386, 265017, 265018, 265019, 264683, 265020, 874684.8, 5442344	65274572	264907, 264909	264488, 35696286, 28331825, 29331828, 264508, 264509, 264906, 264907, 264908, 264901, 264910, 265011, 264760, 264681, 264769, 264689, 21906765, 264693, 264628, 18108370,	264629, 264631, 264634, 264563, 264564, 264566, 264486
UNCLASSIFIED	мнс	ATPase_associated 18108351	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	transcriptfactor	UNCLASSIFIED	UNCLASSIFIED	
							Contains protein domain (PF00170) - transcriptfactor bZIP transcription factor			
2807 91720702 (5613, 5614) Novel Protein sim. GBank gil4468310[emb CAB37991 - (AL031432) dJ465N24.1 (PUTATIVE novel protein similar to predicted yeast and worm proteins) [Homo sapiens]	Novel Protein sim. GBank gi[5541863 emb CAB51071.1 - (AL096857) hypothetical protein [Homo sapiens]	Novel Protein sim. GBank gi 2772561 (AC004002) - similar to ciliary dynein beta heavy chain; 78% Similarity to P23098 (PID:g118965) [Homo sapiens]			88093334 (5623, 5624) Novel Protein sim. GBank gij3264583 (AC005189) - match to ESTs H97756 (NID:g1118643) and AA085546 (NID:g1629773) [Homo sapiens]	91218755 (5625, 5626) Novel Protein sim. GBank gi 4240273 dbj BAA74915.1 - (AB020699) KIAA0892 protein [Homo saplens]	Novel Protein sim. GBank gi[3548791 (AC005620)		9338&£8 (3631, 3632) Novel Protein sim. GBank gi[5420389 emb CAB46680.1 - (AJ243460) proteophosphoglycan [Leishmania major]	
91720702 (5613, 5614)	95359111 (5615, 5616)	88083530 (5617, 5618)	87259032 (5619, 5620)	91235845 (5621, 5622)	88093334 (5623, 5624)	91218755 (5625, 5626)		2815 79774521 (5629, 5630))))	
2807	2808	2809	2810	2811		2813	2814	2815	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	

264259, 29331822, 29331827, 264508, 264509, 264509, 264908, 264908, 264909, 264909, 264909, 264909, 264509, 265593, 264593, 264788, 255006, 264769, 284764, 264288, 264687, 264769, 263978, 264568, 264686, 264568, 264486, 264567	66712502	264909, 264511	18108394, 52646365, 52644045, 264112, 265009, 21906754, 265017, 18108361, 264683, 264369, 264686, 264769, 21906769, 33657023, 33657109, 18108370, 18108374, 18108377, 18108385	264636	264259, 29331822, 60432289, 264905, 60433356, 33657402, 265011, 264760, 21906765, 60170615, 264638	264766	264907	264760	52645156, 22278994, 35696286, 22278997, 22278998, 52645080, 29331822, 29331824, 29331827, 35696052, 29331824, 33656970, 52644046, 52646317, 3365084, 52644296, 265017, 265018, 265019, 265018, 21906764, 21906766, 21906766, 21906767, 21906769, 35695917, 52644150, 33657109, 33657109, 27486265, 35695763, 35696423, 35695845, 52644332, 18108385, 87168518, 264484
	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	complement	UNCLASSIFIED	complement	UNCLASSIFIED		
	Contains protein domain (PF00023) - UNCLASSIFIED Ank repeat		Contains protein domain (PF00919) - UNCLASSIFIED Uncharacterized protein family UPF0004	Contains protein domain (PF00386) - complement C1q domain		Contains protein domain (PF00386) - complement C1q domain			
Novel Protein sim. GBank gij1293846 (U56966) - coded for by C. elegans cDNA yk30b3.5; coded for by C. elegans cDNA yk30b3.3 [Caenorhabditis elegans]	88073579 (5635, 5636) Novel Protein sim. GBank gij549986 (U13149) - possible apospory-associated protein [Pennisetum ciliare]		2820 87765744 (5639, 5640) Novel Protein sim. GBank gil4929773 gb AAD34147.1 AF15209 - (AF152097) CGI-05 Uncharacterized protein family protein [Homo sapiens] UPF0004	95320511 (5641, 5642) Novel Protein sim. GBank gij399144 sp P02747 C1QC_HUMAN - COMPLEMENT C1Q C1q domain SUBCOMPONENT, C CHAIN PRECURSOR	94260221 (5643, 5644) Novel Protein sim. GBank gi[2224671 dbj BAA20820 - (AB002363) KIAA0365 [Homo sapiens]	95320513 (5645, 5646) Novel Protein sim. GBank gij399144 sp P02747 C1QC_HUMAN - COMPLEMENT C1Q C1q domain SUBCOMPONENT, C CHAIN PRECURSOR	95320515 (5647, 5648) Novel Protein sim. GBank gij399144 sp P02747 C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR		94311905 (5651, 5652) Novel Protein sim. GBank gij3859683 emb CAA22020] - (AL033503) conserved hypothetical protein [Candida albicans]
		87793527 (5637, 5638)	87765744 (5639, 5640)	95320511 (5641, 5642)	94260221 (5643, 5644)	95320513 (5645, 5646)	95320515 (5647, 5648)		l l
2817	2818	2819	282(2821	2822	2823	2824	2825	2826

		22278996, 22278997, 264091, 264093, 60432049, 284259, 29331822, 29331825, 29331825, 29331825, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 29331827, 264903, 260433438, 21906764, 265011, 264693, 21906768, 21906768, 21906768, 21906769, 29148629, 52644150, 264693, 33657109, 18108374, 264634, 18108385, 60432113, 22279000, 264565, 264486		FIED 265017 FIED 22278997, 22278999, 52646317, 264288, 264688, 21906767, 60431528, 264638, 22279000		iated	ictor 264259, 264112, 265010, 264762, 264764, 263974, 264555, 264558, 264559
6) - complemen	22) - phosphata			UNCLASSIFIED	UNCLASSIFIED	ATPase_a	12) - transcriptfa
Contains protein domain (PF00386) - complement	Contains protein domain (PF0010 Protein-tyrosine phosphatase						Contains protein domain (PF00412) - transcriptfactor I IM domain containing professis
2827 95320519 (5653, 5654) Novel Protein sim. GBank gij399144tsplP02747/C1QC_HUMAN - COMPLEMENT C1Q C1q domain SUBCOMPONENT, C CHAIN PRECURSOR	91229615 (5655, 5656) Novel Protein sim. CBank gij3598974 (AF077000) - protein Contains protein domain (PF00102) - phosphatase tyrosine phosphatase TD14 [Rattus norvegicus]	87651244 (5657, 5658) Novel Protein sim. GBank gil4680689lgblAAD27734. 1 AF13295 - (AF132959) CGI-25 protein [Homo sapiens]	88087109 (5659, 5660) Novel Protein sim. GBank gi[2498667]sp Q61200 NPH1_MOUSE - NEUREXOPHILIN 1		87612938 (5665, 5666) Novel Protein sim. GBank gi[5262615 emb CAB45747.1 - (AL080156) hypothetical protein [Homo sapiens] 86974703 (5667, 5668) Novel Protein sim. GBank gi[2224567 dbj]BAA20772 -	(10 2/ 1	Novel Protein sim. GBank gi[2351568 (U76618) - N-RAP Miss misculus]
95320519 (5653, 5654)	91229615 (5655, 5656)	87651244 (5657, 5658)	88087109 (5659, 5660)	87614717 (5661, 5662) 87631809 (5663, 5664)	87612938 (5665, 5666) 86974703 (5667, 5668)	87775712 (5669, 5670)	85724748 (5671, 5672) Novel Protein sim. [Mus musculus]
2827	2828	2829		2831	2833		2836

	UNCLASSIFIED 18108394, 18108397, 264259, 29331826, 265007, 265019, 264448, 18108368, 20281149, 264565, 264567	UNCLASSIFIED 264555	264509, 264511, 265011, 264288, 264769,	265020, 264634, 264636, 264556	bind 56182575, 35696052, 264907, 264908,	265022, 264691, 33657182, 35695763,	18108370, 35695855, 264631, 264559, 264563, 264567	264600	UNCLASSIFIED 264906, 264907, 264908, 264909, 264910,	264764, 35695855, 83373044, 18108385	264685		lox 29146498, 87168474, 264686, 35696423.	UNCLASSIFIED 29331830, 264909, 265008, 265011.			UNCLASSIFIED 264490, 264259, 264508, 264905, 264907,	264510, 265007, 265008, 264591, 264592,	264593, 264594, 264595, 55812038, 265010,	263011, 264604, 264703, 264704, 264703, 264766, 264686, 264628, 264629, 264555,	264636, 264556, 264557, 264638, 264558,	264559, 264563, 264566, 264567	29331822, 35696052, 264509, 264906,	265007, 264594, 265018, 264288, 263972,	UNCLASSIFIED 35695286, 265008, 265009, 269018, 26428, 35695917, 264693, 18108374, 35695855,
UNCLA	UNCLA	UNCLA			Contains protein domain (PF00076) - dna_rna_bind	RBD, or RNP domain)			UNCLA		Contains protein domain (PF00008) -	EGF-like domain	нотерох	UNCLA		UNCLA	UNCLA								ONCEA
2837 87766482 (5673, 5674) Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan Leishmania major	87775392 (5675, 5676) Novel Protein sim. GBank gi 973378 (U31263) - core protein [Hepatitis C virus]		2840 87774565 (5679, 5680) Novel Protein sim. GBank gi 1575515 (U64899) -	thrombospondin-related anonymous protein (Plasmodium gallinaceum)	86982568 (5681, 5682) Novel Protein sim. GBank gij2224605 dbj BAA20790 -				91012494 (5685, 5686) Novel Protein sim. GBank gi[5578957]emb CAB51350.1 -	(AL050306) dJ475B7.1 (novel KIAA0615 and KIAA0323 LIKE protein) [Homo sapiens]	56731154 (5687, 5688) Novel Protein sim, GBank	1955-1231sp)Q08878FBLC_MOUSE - FIBULIN-1, ISOFORM C PRECURSOR (BASEMENT-MEMBRANE PROTEIN 90) (8M-90)	94321719 (5689, 5690) Novel Protein sim. GBank gij5420387 emb CAB46679.1 - (AJ243459) proteophosopoolycan (Leishmania maior)	88318613 (5691, 5692) Novel Protein sim. GBank	gi 5306263 gb AAD41995.1 AC00623 - (AC006233) unknown protein [Arabidopsis thaliana]	81811757 (5693, 5694) Novel Protein sim. GBank gij3399676 (AC005390) - R31180_1 [Homo sapiens]	Novel Protein sim. GBank gi 5262615 emb CAB45747.1 -	(AL080155) hypothetical protein [Homo sapiens]					88084283 (5697, 5698) Novel Protein sim. GBank gi 3342218 (AC004131) -	Unknown gene product [Homo sapiens]	
837 87766482 (5673, 5674)	2838 87775392 (5675, 5676)	339 85799317 (5677, 5678)	940 87774665 (5679, 5680)		2841 86982568 (5681, 5682)			2842 80080086 (5683, 5684)			2844 56731154 (5687, 5688)		2845 94321719 (5689, 5690)	2846 88318613 (5691, 5692)		2847 81811757 (5693, 5694)	2848 87612943 (5695, 5696)						2849 88084283 (5697, 5698)		2850 87623636 (5699, 5700)

2851	87820548 (5701 5702)	2851 87820548 (5701 5700) Novel Protein sim CBack all 1223 (5101 510 1)			
		(AF051098) seven transmembrane domain orphan receptor		UNCLASSIFIED	264906, 264907, 56182435, 264758,
		[Mus musculus]			264369, 21906765, 264691, 264692, 264693,
2852	_	Name Booksia aim Const. Handerson Handeson			35695855, 264556, 18108385, 264567
		drosovosa (2, 5, 5, 5, 5, 5) I Nover Frideni sim. Gaank gij 1823/29 (U88308) - similar to drosophila membrane protein PATCHED SP:P18502 (PID:g129645) [Caenorhabditis elegans]		UNCLASSIFIED	264591
2853		87784630 (5705, 5706) Novel Protein sim. GBank gi 2702347 (AF027503) - putative Contains protein domain (PF00397) - kinase	Contains protein domain (PF00397) -	kinase	56182575 55811150 264690 27486262
		membrane-associated guanylate kinase 1 [Mus musculus]	WW domain		27486265, 264632, 56182323, 56526486,
2854		lch Sich	Contains protein domain (PF01344) - dna ma bind	dna rna bind	35696286. 29331824. 29331826. 29331828
			Kelch motif	ı	264908, 264768, 264693, 22279002, 264482
		bridges; 77% Similarity to A45773 (PID:g1079096) [Homo sapiens]			
2855		94723856 (5709, 5710) Novel Protein sim. GBank gij1504040 dbj BAA13219 -	Contains protein domain (PF00560) - glycoprotein	glycoprotein	22278994, 29331822, 29331824, 29331825.
		(U000903) Similar to D. melanogaster peroxidasin(U11052)	Leucine Rich Repeat		264906, 264908, 265009, 33109954, 265018,
					265019, 264448, 21906765, 265020, 264690,
					27486265, 83373044, 22279000, 22279002,
2856	_	88093359 (5711, 5712) Novel Protein sim. GBank gil3264583 (AC005189). match			704407
		to ESTs H97758 (NID:g1118643) and AA085546			21906766, 22278997, 265022, 29331822.
		(NID:g1628773) [Homo sapiens]			28331626, Z/486262, Z650U/, Z65009, 265017 264482 264552 49408254
2857		1855 (AC004537) - similar	Contains protein domain (PE00628) - struct	strict	20278006 3660636 20234834 20234825
			PHD-finger		22270990, 30090200, 29331624, 29331623, 35696052, 264103, 264108, 86182434
		(PID:g2829208) [Homo sapiens]			21906765, 21906769, 265020, 18108368
285R		87434748 (5715, 5716) Navel Bertain sin Co			35695763, 22279002, 264563
3			Contains protein domain (PF00097) - dna_rna_bind		264569, 264887, 22278995, 22278996,
			Zinc finger, C3HC4 type (RING		22278997, 22278999, 264259, 29331826,
		THOLEIN MEL-10 (ZINC FINGER PROTEIN 144)	finger)		29331827, 29331828, 264509, 264905,
					264906, 29331830, 264908, 52644045.
					264909, 264511, 264512, 265007, 265008,
					264910, 265009, 264593, 60433356, 264595,
					264758, 21906754, 265010, 265011, 264604,
					265018, 264760, 18108351, 264763, 264682,
		-			264764, 264765, 264288, 264369, 264685,
					264766, 264768, 18108357, 264769,
	-				21906766, 21906767, 265021, 264534,
					60170615, 264691, 264692, 18108370,
					264629, 18108374, 264631, 264636, 263981,
					18108381, 264558, 18108385, 22279002,
2859		90937675 (5717, 5718) Novel Protein sim. GBank nit4325320 ninhta 017331 11			264564, 264566, 264486, 264567
		(AF124427) claudin-15 [Mus musculus)		UNCLASSIFIED	60424179, 65274572, 29331828, 264905,
					264511, 264758, 265011, 21906767,
					21906769, 55811957, 265021, 56182323

264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264510, 264511, 33109954, 18108351, 264683, 264765, 264369, 264686, 21906765, 264691, 264692, 264693, 18108388, 22279002,	264369, 264692	264691, 264638	56994075, 35696286, 22278998, 29331822, 29331824, 35686052, 29331828, 264106, 264511, 55812038, 33657084, 55811386, 265018, 265019, 21906765, 21906766, 21906769, 3565709, 265020, 265022, 3365709, 33657399, 264622, 18108376, 60431850, 56182323, 18108385, 87168518, 22726002	264259, 264910	22278995, 21906764, 264482	65274572, 22278996, 22278999, 264259, 264259, 29331824, 29331827, 29331828, 60433438, 21906754, 265018, 264448, 264784, 576444150, 83373344	264488, 264768, 21906768, 22278938, 265022, 264259, 264508, 264305, 264907, 264511, 264910, 264635, 264636, 264636, 264537, 265011, 265017, 265018, 265019, 264563, 264688, 264288, 264766, 264369, 264567, 264486, 264288, 264766
UNCLASSIFIED		UNCLASSIFIED	struct	struct		glycoprotein	UNCLASSIFIED
			Contains protein domain (PF00096) - struct Zinc finger, C2H2 type				
2860 87532599 (5719, 5720) Novel Protein sim. GBank gil4469186[emb CAB38414.1]. (AL031589) d.1163J1.2.1 (novel protein similar to C. elegans B0035.16 and bacterial tRNA (5- Methylaminomethyl-2-thiouridylate)-Methyltransferases) (isoform 1) [Homo sapiens]	86698507 (5721, 5722) Novel Protein sim. GBank gij3941730 (AF108083) - BS4 [Homo sapiens]	87569585 (5723, 5724) Novel Protein sim. GBank gil4505013 ref NP_002310.1 pLRN - leucine-rich neuronal protein	91220421 (5725, 5726) Novei Protein sim. GBank gij3249088 (AC004473) - Contains similarity to goliath protein gb M97204 from D. melanogster. [Arabidopsis thaliana]	87420030 (5727, 5728) Novel Protein sim. GBank giļ1079451 pir 455463 - tropomodulīn. skeletal muscle - chicken	95312191 (5729, 5730) Novel Protein sim. GBank gil438840 (L19048) - MSA-2 [[Plasmodium falciparum]	95105480 (5731, 5732) Novel Protein sim. GBank gi 585703 sp Q07066 PMP2_RAT	86908001 (5733, 5734) Novel Protein sim. GBank gil4580997Igb AAD24571.1 AF12108 - (AF121081) cAMP inducible 2 protein [Mus musculus]
87532599 (5719, 5720)	86698507 (5721, 5722)	87569585 (5723, 5724)					86908001 (5733, 5734)
2860	2861						2867

18108392, 264488, 22278994, 22278997, 22278998, 22278999, 60432049, 264259, 29331824, 29331825, 60432289, 29331824, 29331825, 60432289, 29331827, 35696052, 29331828, 264508, 265009, 265009, 265009, 264591, 6043229, 264597, 6043356, 264597, 6043229, 264597, 6043256, 204597, 6043258, 265017, 265019, 264682, 264591, 6043268, 265021, 265022, 264691, 33657182, 18108368, 27486264, 27486264, 27486264, 27486264, 27486264, 22279000, 22279000, 22279000, 22279000, 22279000, 22279000, 22279000, 22279000, 22279000, 224887		IFIED 264905, 264908, 264764, 21906769, 264634	FIED 264259, 29331822, 60432289, 29331827,	264764, 18108354, 265021, 27486265.	263981		204753, 18108370, 204907, 204906, 204906, 204906, 264764, 264267, 264909, 264486, 264766, 18108391
	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED		kinase	UNCLASSIFIED	
2868 95303283 (5735, 5736) Novel Protein sim. CBank gi 1292868 emb CAA63923 - (X94232) t-Cell activation protein [Homo sapiens]			2871 88318621 (5741, 5742) Novel Protein sim. GBank	unknown protein [Arabidopsis thaliana]	95312197 (5743, 5744) Novet Protein sim. GBank gil112205 pir B39066 - proline- nich protein 15 - rat		
9530328 3 (5735, 5736)	88094412 (5737, 5738)	84404574 (5739, 5740)	88318621 (5741, 5742)		95312197 (5743, 5744)	88094252 (5745, 5746)	
2868	2869	2870	2871		2872	2873	

52644507, 52845156, 52846365, 56182575, 22278994, 22278995, 56994075, 35696286, 60432049, 264259, 56994075, 35696286, 60432049, 264259, 52645080, 29331822, 56182181, 29331824, 60424269, 29331827, 3696052, 29331828, 33656970, 66712502, 52644045, 255007, 285009, 66433356, 264758, 55812038, 18108348, 52846317, 33109954, 33657084, 265017, 264604, 265018, 265019, 264682, 264369, 26504150, 265020, 265021, 265022, 52644150, 33657023, 52645129, 18108374, 18108376, 3569423, 56182323, 18108374, 18108376, 3569423, 56182323, 18108378, 264563, 264565	22278996, 22278997, 22278999, 29331826, 29331828, 29331828, 29146499, 66712502, 265008, 265017, 18108351, 264683, 264689, 21906767, 18108376, 18108377, 55811576, 60170394, 22279000, 264487		52646365, 65274572, 56182575, 22278997, 22278998, 22278999, 60432049, 52645080, 60424269, 60432289, 29331827, 35696052, 29331828, 66712502, 52644046, 56182435, 60433356, 33657402, 33657084, 265019, 55811150, 264448, 264369, 21906769, 25906769, 25906769, 25906769, 2564332, 60431528, 55811576, 52644332, 56526486, 60432113, 22279000, 22279002, 264563	264488, 264259, 29331822, 29331826, 264905, 264509, 264907, 265006, 264509, 264907, 264909, 264750, 265006, 264511, 264512, 33657402, 264768, 264288, 264684, 264685, 264768, 264689, 264690, 264699, 264699, 264629, 264630, 264630, 264630, 264630, 264630, 264630, 264566, 264566, 264563, 264567, 264563, 264564, 264565, 264566, 264486, 264567, 264567
UNCLASSIFIED 5264456 2227895 6043204 661441 8671441 2933187 265020 336570 2190670 2190670 2190670 3365020 336570 336570 264682 265020 336570 336570 336570 336570 264682 264682 264682		fgf		
	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat	Contains protein domain (PF00167) - fgf Fibroblast growth factor		Contains protein domain (PF01207) - UNCLASSIFIED Uncharacterized protein family UPF0034
2874 94313549 (5747, 5748) Novel Protein sim. GBank gij3212854 (AC004005) - unknown protein [Arabidopsis thaliana]	88083726 (5749, 5750) Novel Protein sim. GBank gi 2781386 (AC004010) - similar to Leucine-rich transmembrane proteins; 44% similarity to U42767 (PID:g1736918) [Homo sapiens]	88090854 (5751, 5752) Novel Protein sim. GBank gil2979530 (AC004449) - R33683_2 [Homo sapiens]	94747029 (5753, 5754) Novel Protein sim. GBank gil4704208lęmb CAB41646.1 - (AL035419) dJ1100H13.1 (putative novel protein) [Homo sapiens]	88095309 (5755, 5756) Novel Protein sim. GBank gij3876775jembjCAB03067] - (Z81077) predicited using Genefinder; Similarity to Yeast protein 8248 (TR:G587531) [Caenorhabditis elegans]
574 94313549 (5747, 5748) N		2876 88090854 (5751, 5752)	2877 94747029 (5753, 5754)	2878 88095309 (5755, 5756)

2870	87860122 (5757 575R)	2870 87860122 (5757 5758) Noval Protein sim CBank alusos145 and 20752 11	Cestaine arabain demain (BE00315) 11NICLASSIETED	ODICION IONI	10400250 DE4050 DE400E 40400970
	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	(AF127374) unknown (Streptomyces lavendulae)	Uracil-DNA divcosvlase		264629 264908 264909 18108374
	_				18108377 265006 264910 264637
					60170394, 264559, 265017, 264564, 264565.
					264567, 264684, 264369
2880	94851439 (5759, 5760) Novel Protein sim.	Novel Protein sim. GBank			264488, 52646365, 52646842, 22278994,
		gil4680703 gb AAD27741.1 AF13296 - (AF132966) CGI-32			35696286, 22278998, 22278999, 264259,
		protein [Homo sapiens]			29331822, 29331824, 29331825, 29331826,
					29331827, 35696052, 29331828, 264107,
_					264508, 264509, 264905, 264906, 264907,
					264908, 264909, 52644045, 264510, 265006,
					264511, 265007, 264512, 265009, 264910,
					264594, 21906754, 52646317, 52644296,
					87168559, 264600, 264604, 264605, 264760,
					264764, 264288, 264766, 264768, 264687,
					264769, 21906766, 21906769, 35695917,
	_				265021, 264690, 264692, 33657023,
					52645129, 33657109, 33657182, 27486262,
					33657349, 264629, 18108374, 35695855,
					264634, 264635, 264636, 264637, 264638,
					264557, 52644332, 264558, 264559
					83373044 384404 323750000 384863
					033/3044, Z04404, ZZZ/8000, Z04303, 264483 264567 264486
2881	87650539 (5761 5762)	87650539 (5761 5762) Novel Protein sim GBank oil733571 (103452) . No		INCI ASSIEIED	22278008 20331822 52544045 21906765
<u> </u>	(2010,1010)	definition line found (Caenorhahditis elegans)		מונים מונים מונים	222/0930, 2333/022, 32044043, 2 (300/03, 254639, 6043213
Caac	7	Mound Destation of the Control of th			2017000 0001000 0001000 0001000
7007		Novel Protein Sim. GBank gill118112 (U41559) - No			264488, 22278996, 22278999, 29331822,
		definition line found [Caenorhabditis elegans]			29331826, 264908, 60170831, 60433356,
					55812038, 264681, 264682, 264686, 264687,
					264688, 21906768, 21906769, 264693,
					263967, 18108374, 55811576, 56182323,
	_				22279002, 264566
2883	95362875 (5765, 5766) Novel Protein sim.	Novel Protein sim. GBank	Contains protein domain (PF00400) - UNCLASSIFIED	UNCLASSIFIED	264112, 264682
		igi 4868008 gb AAD31087.1 AF10693 - (AF106934) vitamin Direcentorinteracting protein (Homo canions)	WD domain, G-beta repeat		
2884		ferradas official transit formation of the state of the s			265018, 264634
2885	83006306 (5769, 5770) Novel Protein sim.	Novel Protein sim. GBank gi[2224697 db BAA20832 -		UNCLASSIFIED	264686, 264693
	_	(AB002376) KIAA0378 [Homo sapiens]			
2886		91237823 (5771, 5772) Novel Protein sim. GBank gij1255889 (U53344) - T07H6.5	Contains protein domain (PF00084) - complementrecept	1	60432049, 264259, 29331828, 264908,
		gene product [Caenorhabditis elegans]	Sushi domain (SCR repeat)		264511, 264595, 60433438, 264596, 265017,
					264605, 263969, 263972, 264555, 83373044,
	_				87168518, 264566
2887		91227860 (5773, 5774) Novel Protein sim. GBank gij3882323 dbj BAA34521.1 -		UNCLASSIFIED	18108351, 264686, 264629, 264631, 264639,
		(AB018344) KIAA0801 protein [Homo sapiens]			83373044, 264482
2888			Contains protein domain (PF00096) - transcriptfactor	transcriptfactor	264488, 264259, 29331828, 264508, 264906.
		gi 4508015 ref NP_003447.1 pZNF2 - zinc finger protein 205 Zinc finger, C2H2 type	Zinc finger, C2H2 type		264593, 264758, 264766, 264769, 18108374,
					83373044, 264486

2889	87606562 (5777, 5778)			UNCLASSIFIED	56994075, 29331824, 265009, 264760, 18108354, 264288
2890	78703853 (5779, 5780)	78703853 (5779, 5780) Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6]		UNCLASSIFIED	264591, 264766
2891		Novel Protein sim. GBank gil3877750 emb CAB01508 - (Z78064) predicted using Genefinder; similar to collagen; cDNA EST EMBL:D65865 comes from this gene; cDNA EST EMBL:D69651 comes from this gene; cDNA EST EMBL:D66026 comes from this gene; cDNA EST EMBL:D69658 comes from this gen	•	UNCLASSIFIED	264591, 264593, 264369, 264685, 264693, 264628, 264563, 264566
2892		95419745 (5783, 5784) Novel Protein sim. GBank gil4929759/gb AAD34140.1 AF15190 - (AF151903) CGI-145 protein [Homo sapiens]		UNCLASSIFIED	65274572, 35696286, 29331828, 264110, 265009, 60433438, 265018, 265019, 18108351, 264288, 21906765, 21906766, 21906769, 55811957, 264690, 65274620, 263967, 3569585
2893				UNCLASSIFIED	264259, 265010, 264682, 18108370, 264555, 264556
2894		87755985 (5787, 5788) Novel Protein sim. GBank gi 5669015 gb AAD46135.1 - (AF080171) zinc finger protein ZNF232 [Homo sapiens]	Contains protein domain (PF00096) - Iranscriptfactor Zinc finger, C2H2 type	transcriptfactor	264259, 265006, 60433438, 52644296, 265011, 264369, 35695917, 18108381, 18108382, 18108388
2895		86938778 (5789, 5790) Novel Protein sim. GBank gil3924708 emb CAA84646 - (235597) Weak similarity with sea squirt nidogen precursor protein (blastp score 71); cDNA EST EMBL:102069 comes from this gene; cDNA EST EMBL:D73147 comes from this gene;	Contains protein domain (PF01437) - Plexin repeat		29331824, 265007, 264762, 264636, 264563
2896		87752122 (5791, 5792) Novel Protein sim. GBank gil4885549 ref NP_005456.1 pPKBG - protein kinase B gamma	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	kinase	18108392, 18108394, 18108398, 22278998, 264259, 264259, 29331822, 29331824, 29146499, 264908, 265007, 265009, 265018, 265019, 264689, 264689, 264689, 264689, 264639, 264639, 18108384, 18108384, 264567
2897	96413057 (5793, 5794)	95413057 (5793, 5794) Novel Protein sim. GBank gil4502877 ref\nP_001296.1 pCLDN - Clostridium perfringens enterotoxin receptor 1		UNCLASSIFIED	60424179, 56182575, 22278996, 35696286, 22278997, 22278999, 60432049, 264259, 29331824, 29331828, 560424269, 60432289, 29331824, 29331828, 35696052, 264908, 56182435, 256009, 264910, 60170831, 60431735, 60433356, 60433438, 65274444, 55811386, 256018, 18108351, 264448, 264288, 264687, 21906765, 21906767, 21906769, 55811957, 35695917, 26428, 264629, 60431528, 18108334, 255810764, 55811576, 35696333, 264528, 264628, 264639, 60431528, 18108334, 255810764, 55811576, 35696385, 264555, 56182323, 18108385, 264404, 22279000, 22279000, 264556, 56182323, 18108385, 264404,
2898	87750340 (5795, 5796)			UNCLASSIFIED	22278995, 264604, 18108385, 264566

- 1	_	_	_	_	_					_		_									_				_			_	_			1			_	-		_
	264764, 21906764, 264692	65274572, 56182575, 35696286, 60432049,	264259, 29331824, 66714117, 29331826,	35696052, 29331828, 66712502, 56182435,	265006, 265007, 265008, 265009, 60433356,	264758, 265018, 264764, 264765, 264288,	264768, 21906764, 21906768, 21906769,	265020 264602 264603 32833096 264634	203020, 204032, 204033, 32633386, 204031,	833/3044, 60432113	22278995, 22278997, 22278999, 60432049,	29331822, 29331824, 29331825, 29331827,	35696052, 33656970, 264910, 265009,	21906754, 33657084, 87168474, 265010,	265018, 21906764, 21906765, 21906766,	21906767, 21906769, 33657023, 264693,	33657109, 33657349, 35696423, 35695855,	263981, 56182323, 22279002	264683, 264691		264887, 22278994, 264259, 29331826	29331828, 264905, 52644045, 56182435,	264511, 265017, 265018, 18108351, 264448,	264683, 264769, 264689, 35695917,	52644150, 87168518, 60432113, 22279002	52646365, 22278999, 264259, 35696052,	265011, 265017, 264683, 264769, 35695917,	265020, 263967, 18108374, 35695855,	264637, 264952, 18108385, 18108387	Z65009, Z64681, Z6468Z		22278996, 35696286, 22278998, 264259,	60432289, 29331828, 29331830, 66712502,	265009, 60170831, 33109954, 264448,	264683, 264288, 264689, 21906766,	21906767, 21906768, 55811957, 35695917,	265022, 52644150, 264691, 33657023,	264692, 264693, 35695855, 60432113,
	UNCLASSIFIED	glycoprotein									phosphatase															UNCLASSIFIED				oncogene		UNCLASSIFIED						
																																	PX domain					
		Novel Protein sim. GBank gil4581470 emb CAB40137.11-	(Y18483) SLC/A8 protein [Homo sapiens]								or444/31 (3001, 3002) Novel Protein Sim. GBank	gil4759272 ref NP_004614.1 pTTC4 - tetratricopeptide	repeat domain 4						85745271 (5803, 5804) Novel Protein sim. GBank gi[2414615 emb CAB16364] -	(zeazes) hypometical protein (scrizosaccharomyces pombe)	87606733 (5805, 5806) Novel Protein sim. GBank gi 1079318 pir S52241 - XLCL2	protein - African clawed frog				86458072 (5807, 5808) Novel Protein sim. GBank	gij5639823lgb[AAD45885.1 AF14367 - (AF143676)	multispanning nuclear envelope membrane protein nurim	84440006 (5800 5810) Novel Dratain sim CBook	oit28837Isot939194IALIT HIMAN : IIII ALLI SUBEAMILY	SQ WARNING ENTRY IIII	95341051 (5811, 5812) Novel Protein sim. GBank	gil4689256 gb AAD27831.1 AF12185 - (AF121858) sorting	nexin 8 [Homo sapiens]				
10000 -0000	Z088 (0030/0/0 (0/8/: 0/88)	94233538 (5789, 5800)								┰	0/444/31 (3001, 3602)					-		_								86458072 (5807, 5808)		<u> </u>	_	מיניים (ספסי ספריים)		95341051 (5811, 5812)						
3	Reo 7	2800		_						200	067								2905		2903					2904			2005			2906						

2807	2907 91211383 (5813, 5814) Novel Protein sim. G	Bank gij1707079 (U80451) - contains	Contains protein domain (PF00226) - eph	eph	52644507, 56182575, 56181686, 22278995, 5604075, 35606786, 60432040, 56182181
		[Caenorhabditis elegans]			35696052, 60431735, 264595, 55812038,
					21906754, 55811386, 265019, 264682,
					264369, 56181562, 21906766, 55811957,
					35695917, 265020, 265021, 33657023,
					33657109, 60431528, 55811576, 35696423,
					35695855, 264638, 22279000
2908	80414246 (5815, 5816) Novel Protein sim.	Novel Protein sim. GBank gij2673917 (AC002561) - putative		helicase	265009, 33109954, 18108351, 264766,
		ATP-dependent RNA helicase [Arabidopsis thaliana]			265021, 264691, 264692, 18108374, 264556,
_					264638, 264557, 264558
5808	87420225 (5817, 5818)			ydə	264259, 87168474, 265018, 18108365,
		- 1			264628
2910	86601075 (5819, 5820) Novel Protein sim.	9			22278995, 264509, 264512, 265007,
		(AL035539) putative protein [Arabidopsis thaliana]			33657402, 265017, 264369, 265022,
					18108365, 264628
2911	94216615 (5821, 5822)	94216615 (5821, 5822) Novel Protein sim. GBank gil4469187 emb CAB38415.11 -		glucoamylase	52646365, 18108397, 22278995, 22278997,
		(AL031588) dJ1163J1.3 (novel protein similar to mouse			22278998, 22278999, 29331824, 29331825,
		899) [Homo sapiens]			52644045, 265006, 265018, 264448,
_					21906765, 21906766, 21906767, 21906768,
					21906769, 55811957, 265021, 18108370,
					18108372, 18108374, 22279000
2912	87731803 (5823, 5824)	Novel Protein sim. GBank	Contains protein domain (PF00904) -		52645156, 264092, 60432049, 264259,
		gi 4929637 gb AAD34079.1 AF15184 - (AF151842) CGI-84 Involucrin repeal	Involucrin repeat		52645080, 29331824, 29331825, 66712502,
		protein [Homo sapiens]			33109954, 264760, 264683, 264288, 264686,
					265021, 264693, 18108368, 263976, 264404
2913	87713823 (5825, 5826)	87713823 (5825, 5826) Novel Protein sim. GBank gil854065 emb CAA58337 -	Contains protein domain (PF00096) - UNCLASSIFIED	UNCLASSIFIED	52644507, 52645156, 52646842, 56182575,
		(X83413) U88 [Human herpesvirus 6]	Zinc finger, C2H2 type		35696286, 22278997, 264259, 52645080,
					29331827, 35696052, 29331828, 264828,
					52644045, 56182435, 55812038, 52646317,
					21906754, 52644296, 87168474, 265017,
					265018, 265019, 18108351, 264682, 264686,
					264689, 21906765, 21906766, 21906767,
					21906768, 35695917, 265020, 52644150,
					27486261, 27486262, 27486265, 35695763,
					55811576, 35695855, 52644332, 22279000,
					22279002, 264563
2914	87797300 (5827, 5828)				264557

2018	188081073 (F820 F820)	2015 188081072 (F820) Keach Named Britain aim Count	10300030		COLUMN COLLAGO POOL COLLAGO COLLAGO
}	(222) = 1222	ail5174485IrefiNP 006030.1lpKIAA - endocytic receptor	Lectin C-type domain	<u>.</u>	204303, 204400, 204001, 204700, 21900700, 52646842 21906767 21906768 56182575
_		(macrophage mannose receptor family)			29148629, 35695917, 22278996, 22278997
		-			22278998, 265021, 22278999, 52644150,
					264691, 264259, 60432049, 264692,
					52645129, 33657109, 33657182, 29331827,
					27486261, 35696052, 29331828, 27486262,
					27486264, 27486265, 33657349, 29146498,
					29146499, 264906, 264907, 18108370,
					264908, 18108372, 52644045, 18108374,
					56182435, 35695855, 264112, 264510,
					265008, 60432229, 264593, 60433356,
_					56182323, 18108382, 55812038, 18108385,
					33109954, 21906754, 33657084, 87168518,
					265010, 265011, 60432113, 265017, 265018,
	_				22279000, 265019, 55811150, 264681,
					18108351, 264763, 264448, 264683, 264566,
2916		95337790 (5831, 5832) Novel Protein sim, GBank qil51048511dbilBAA80165.11		dehydrogenase	18108334, 204308, 204200, 204700 52645156, 65274572, 22278994, 22278995
		(AP000061) 305aa long hypothetical dTDP-4-		,	35696286, 22278996, 22278997, 22278998,
		dehydrorhamnose reductase [Aeropyrum pernix]			22278999, 264259, 29331822, 29331824,
					29331825, 66714117, 60432289, 29331826,
					29331827, 29331828, 33656970, 264509,
					264906, 29331830, 52644045, 264909,
					56182435, 60170831, 264592, 264593,
					33657402, 60433356, 52646317, 21906754,
					33109954, 33657084, 52644296, 85658542,
					265011, 265017, 265018, 265019, 18108351,
					264448, 264288, 52644229, 21906765,
					21906766, 21906767, 21906768, 21906769,
_					55811957, 35695917, 265021, 265022,
					52644150, 33657023, 33657109, 33657182,
					2/486261, 2/486262, 2/486264, 35695/63,
					16106376, 55811576, 55686423, 65274791,
					35695855, 52644332, 264557, 264638,
					56162323, 16106367, 87166318, 22279002, 1264482
2917	87454546 (5833, 5834)	87454546 (5833, 5834) Novel Protein sim. GBank gil3169065 emb CAA19260.1 -		UNCLASSIFIED	60433438, 264602, 264682, 87168518,
		(AL023704) putative transfocation elongation factor-Tu fa			60432113
1	_	mily [Schizosaccharomyces pombe]			
2918	_	85690529 (5835, 5836) Novel Protein sim. GBank gij539218 pir S38038 -			264638
		hypothetical protein YKL201c - yeast (Saccharomyces cerevisiae)			
2919	_	87641497 (5837, 5838) Novel Protein sim. GBank gil2564955 (AF030001) -			66714117, 66712502, 263981
		unknown [Mus musculus]			

2928	95343003 (5851, 5852)				29331828 265011 264768 264680
2927		80408018 (5853, 5854) Novel Protein sim. GBank gi[283032 pitr S22456 - hydroxyntoline-rich alveonolein - perennial taosiota			264764, 264288, 264630, 264637
2928	20452179 (5855, 5856)	+=		INCLASSIFIED	264 650
2020		Novel Dratein cim Coart ailo4433300 File & Angeler		T	800407
6767		(AJ006215) CMP-N-acetylneuraminic acid synthetase [Mus musculus]		UNCLASSIFIED	264569, 264489, 22278994, 35696286, 22278996, 22278998, 22278999, 264094, 264259, 52645080, 29334822, 29334824
					66714117, 29331825, 29331826, 29331827,
					35696052, 33656970, 264109, 29331830,
					52644045, 265009, 33109954, 52644296,
					87168559, 264760, 264762, 264448, 264764, Dexago pexice pexico arcoerce
					204286, 204700, 204768, 21906765,
					21906/66, 21906/68, 21906/69, 35695917, 264691 33657023 264603 33657100
					18108374, 263976, 35696423, 35695855
					263981, 22279000, 22279002, 264567,
2930	95302755 (5859, 5860)			OBIGIOS IONI	20448b
					2010£313,30101000,33030£00,£2£18330,
					FILT. 3000, EET 3000, EG EGG, EGG 10EG,
					56182435, 255010±0, £01505, 02014040, 56182435, 265009, 60170831, 264592
					60432229 60433356 87168474 265010
					265011 265017 265018 265019 264762
					264448, 264683, 264288, 264766, 21906765
					21906769, 35695917, 60170615, 33657023
					33657109, 264628, 18108370, 18108372,
					35696423, 35695855, 264556, 56182323,
					60432113, 264567
2931	_	94312693 (5861, 5862) Novel Protein sim. GBank gi[3786433 (AF098505) - similar C	Contains protein domain (PF00471) - UNCLASSIFIED		52645156, 22278997, 22278998, 29331822,
		to Arabidopsis thaliana male sterility protein 2 (SW:Q08891) Ribosomal protein L33	Ribosomal protein L33		52645080, 29331824, 60432289, 33656970,
		[Caenorhabditis elegans]			60433356, 60433438, 33109954, 21906765,
					21906766, 21906767, 21906768, 265020,
					52644150, 33657023, 33657109, 33657182,
					27486265, 35696423, 35695855, 264555,
2032	79632623 /5863 5864)				8/168518, 60432113, 264566
2023	13022023 (3003, 3004)	No. of Description Of the Control of			264906, 264907
2	(2000, 2002, 2000)	Novel Plotein sim. GBank gij3378056 (AFU1777) - helicase		helicase	264488, 18108392, 56182575, 22278999,
					264091, 264259, 29331825, 60432289,
					29331827, 264508, 52644045, 56182435,
					265007, 265009, 264592, 60433356,
					60433438, 21906754, 265017, 264682,
					264288, 52644229, 21906765, 21906766,
					21906768, 21906769, 265022, 52644150,
					33657023, 33657109, 27486265, 264635,
					264636, 60170394, 56182323, 18108385,
					60432113, 264565, 264566, 264567

22278997, 22278999, 29331824, 33657402, 264991, 27486262, 264628, 87168518, 22279000	56182575, 22278995, 60433356, 33657402, 264758, 33109954, 21906754, 265018, 265019, 264448, 264769, 21906764, 21906765, 265021, 264692, 33657023, 33657109, 33657349, 55810764, 22279000	22278997, 29331827, 29331828, 265009, 265017, 264605, 265020, 55811576, 1810838, 64432113, 264681	56994075, 22278999, 264229, 60432049, 29331822, 56182181, 29331827, 29331828, 264906, 264908, 264909, 56182435, 265006, 264512, 264910, 60170831, 60433356, 265011, 265018, 18108351, 264448, 264288, 264766, 52644229, 21906765, 29148784, 65274791, 264556, 56182323, 60170394,	264488, 56182575, 22278996, 35696286, 22278997, 22278998, 22278999, 264490, 264259, 29331822, 29331824, 66714117, 29331827, 35696052, 264107, 264905, 26572602, 5564404, 56182403, 265708, 56182435, 26411, 265008, 265009, 6043229, 33657402, 60433438, 55812038, 21906754, 8555842, 256010, 265011, 27186559, 265017, 265018, 264681, 264689, 21906765, 21906767, 21906768, 55811957, 35695917, 265020, 60170615, 264690, 264691, 264692, 33657023, 264693, 256988, 264691, 264692, 33657023, 264693, 256988, 264691, 264692, 33657023, 264693, 256988, 264691, 264691, 264692, 33657023, 264693, 256988555, 264565, 264565, 264567, 264691, 264691, 56182323, 60170394, 8373044, 18108385, 56576486, 6043713, 27278900	Contains protein domain (PF00400) - ATPase_associated 264907, 265018, 264681, 264685, 264686 WD domain, G-beta repeat
	UNCLASSIFIED	nuclease	UNCLASSIFIED	ubiquitin	ATPase_associate
		Contains protein domain (PF00856) - nuclease SET domain		Contains protein domain (PF00179) - ubiquitin Ubiquitin-conjugating enzyme	Contains protein domain (PF00400) - WD domain, G-beta repeat
		Novel Protein sim. GBank gil4153862 (AC005065) - determined by GENSCAN prediction and spliced EST; match to EST R84329 (NID:942735) [Homo sapiens]	Novel Protein sim. GBank gi[5174409 ref NP_006101.1 pCD2B - CD2 antigen (cytoplasmic tail)-binding protein 2	Novel Protein sim. GBank gij3319990 emb CAA76720 - (Y17267) ubiquitin-conjugating enzyme [Mus musculus]	Novet Protein sim. GBank gij3979900jembjCAA99909j (Z75547) similar to WD domain, G-beta repeat; cDNA EST yk371b7.5 comes from this gene; cDNA EST yk312h1.5 comes from this gene; cDNA EST yk465d5.5 comes from this gene; cDNA EST yk472c4.5 comes from this gene; cDNA EST yk29218
		87605863 (5871, 5872) Novel Protein sim. determined by GEN match to EST R843	94853096 (5873, 5874) Novel Protein sim. gij5174409 ref NP_(cytoplasmic tail)-bi	95419773 (5675, 5876) Novel Protein sim. (Y17267) ubiquitin-	2939 87786622 (5877, 5878) Novel Protein sim. (275547) similar to (275547) similar to yk371b7.5 comes fi comes from this getting the protein sim. (2500 EST yk29218)
2934	2935	2936	2937	2938	2939

22278996, 29331822, 29331824, 66714117, 29331826, 29331828, 264905, 264908, 66712502, 29331830, 265011, 265017, 264764, 264369, 21906766, 21906767, 33657023, 33657109, 32833986, 18108374, 18108387, 264566	264557	264910, 265010, 264768			60424179, 22278995, 22278996, 22278998, 22278999, 264259, 56182181, 29331824, 60424269, 60432289, 35696052, 264908, 265006, 60433356, 58182039, 264599, 55811386, 265018, 264683, 264684, 18108351, 264448, 264683, 264639, 264599, 264593, 2659591, 264639, 60170394, 83373044, 22279000, 264566, 264567
UNCLASSIFIED	UNCLASSIFIED	glycoprotein	collagen	ATPase_associat	UNCLASSIFIED
				Contains protein domain (PF00415) - ATPase_associated Regulator of chromosome condensation (RCC1)	Contains protein domain (PF00096) - UNCLASSIFIED Zinc finger, C2H2 type
	Novel Protein sim. GBank gi 3413872 dbj BAA32300 - (AB007924) KIAA0455 protein [Homo sapiens]	87430203 (5883, 5884) Novel Protein sim. GBank gi 1172845 sp P46629 RB25_RABIT - RAS-RELATED PROTEIN RAB-25	95314504 (5885, 5886) Novel Protein sim. GBank gil4929653 gb AAD34087.1 AF15185 - (AF151850) CGI-92 protein [Homo sapiens]	95081063 (5887, 5888) Novel Protein sim. GBank gil4678282 emb CAB41190.1 - (AL049660) 1-acylcerol-3-phosphate acyltransferase-like protein [Arabidopsis thatiana]	94233560 (5889, 5890) Novel Protein sim. GBank gil7288318pp39188JaLU1_HUMAN - !!!! ALU SUBFAMILY Zinc finger, C2H2 type J WARNING ENTRY !!!!
2940 95011103 (5879, 5880)	21423370 (5881, 5882) Novel Protein sim. G (AB007924) KIAA045	87430203 (5883, 5884)	95314504 (5885, 5886)		94233560 (5889, 5890)
2940	2941	2942	2943	2944	2945

				,	_
264488, 264259, 264508, 264509, 264906, 264907, 264907, 264910, 264510, 264511, 265007, 264512, 264913, 264512, 264913, 264512, 264913, 264512, 264913, 264694, 264694, 264694, 264629, 264692, 33657109, 264634, 264636, 264636, 264637, 264565, 264566, 264486, 264567	22278995, 22278996, 22278997, 22278999, 29146498, 264508, 2931830, 265007, 265008, 265009, 60432229, 21906754, 265010, 265017, 265019, 264509, 264685, 21906766, 21906766, 21906768, 21906769, 264628, 18108370, 264629, 264630, 18108387, 60432113	52646842, 22278995, 264259, 29331824, 29331825, 29331827, 29331830, 264909, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264686, 21906767, 264691, 264692, 264692, 264692, 264692, 264692, 264667, 264667	22278996, 22278997, 60432289, 29331826, 29331827, 29331828, 35696052, 29146499, 264104, 264107, 264905, 66172502, 244908, 60433356, 60433438, 87168569, 264764, 5264429, 56181562, 21906767, 21906768, 25906769, 265022, 60170615, 33657023, 22279002	264488, 35696286, 264259, 35696052, 264907, 265007, 264910, 265017, 265018, 264288, 264768, 35695917, 265020, 18108362, 18108370, 18108379, 35696423, 65274791, 35695855, 264556, 56526486, 264486	264693
UNCLASSIFIED	UNCLASSIFIED		UNCLASSIFIED	peptidase	
				Contains protein domain (PF00883) - peptidase Cytosol aminopeptidase family	
Novel Protein sim. GBank gil5441952gblAAD43195.1lAF07286 - (AF072864) peroxisomal membrane protein PMP 24 [Homo sapiens]	87362952 (5893, 5894) Novel Protein sim. GBank gij3540281fgb AAC34383.1 - (AF056116) All-1 related protein [Fugu rubripes]	87626527 (5895, 5896) Novel Protein sim. GBank gil5566614[gb[AAB65654.2] - (AF001533) mitogen-induced [Mus musculus]	88175545 (5897, 5898) Novel Protein sim. GBank gil2132923 pir S67133 - probable membrane protein YOR240w - yeast (Saccharomyces cerevisiae)	95086870 (5899, 5900) Novel Protein sim. GBank gil466102[spIP34629 YOJ6_CAEEL - PUTATIVE AMINOPEPTIDASE ZK353.6 IN CHROMOSOME III	87392357 (5901, 5902) Novel Protein sim. GBank gil4688902 emb CAB41450.1 - [(AJ238248) centaurin beta2 [Homo sapiens]
94317315 (5891, 5892)					
2946	2947	2948	2949	2950	2951

2952	95329952 (5903, 5904)	2952 95329952 (5903, 5904) Novel Protein sim. GBank gil5596693 emb CAB51405.1 - (AL096881) hypothetical protein [Homo sapiens]	Contains protein domain (PF00650) - Iranscriptfactor	transcriptfactor	264687, 52645156, 21906766, 21906769, 22278996, 265020, 264690, 60432049, 264259, 264693, 29331822, 18108365, 29331825, 61432289, 33657109, 18108368, 29331825, 50432289, 33657109, 18108368, 284905, 20281149, 264906, 264907, 264905, 264906, 264910, 264635, 264511, 265008, 265009, 264910, 264635, 264638, 60433356, 264638, 60433356, 26279000, 22279000, 264760, 264760, 264563, 26448, 264288
2953	88093575 (5905, 5906)		Contains protein domain (PF00266) - UNCLASSIFIED Aminofransferases class-V	UNCLASSIFIED	18108396, 56994075, 22278996, 29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264907, 264907, 264510, 26451, 264504, 33657402, 264595, 264596, 2644296, 87168559, 264600, 264760, 2644296, 87168559, 264600, 264760, 264681, 18108351, 264764, 264569, 264687, 21906765, 21906765, 21906765, 21906765, 21906766, 2190678, 21906766, 2190678, 21906769, 35695917, 33657349, 18108364, 52645129, 33657349, 264638, 87168518, 264482, 264563, 264565
			Contains protein domain (PF00019) - I Transforming growth factor beta like domain	tgf	29331822
2955		Novel Protein sim. GBank gil3452473 (AF084205) - serine/Ihreonine protein kinase TAO1 [Rattus norvegicus]	Contains protein domain (PF00069) - kinase Eukaryolic protein kinase domain	kinase	264259, 29331822, 29331824, 29331825, 29331826, 35696052, 264908, 52644045, 264512, 60432229, 265018, 265019, 55811150, 264769, 21906767, 21906768, 21906769, 265021, 60170615, 55810764, 264567
		85789745 (5911, 5912) Novel Protein sim. GBank gil4689254[gb AAD27830.1 AF12185 - (AF121857) sorting nexin 7 [Homo sapiens]	Contains protein domain (PF00787) - PX domain		22278996, 264259, 29331827, 264908, 21906768
	90933301 (5913, 5914)	Novel Protein sim. GBank gil4503023 ref NP_000089.1 pCPT2 - carnitine palmitoyitransferase II precursor		cadherin	22278999, 264259, 29331824, 29331827, 265008, 264595, 264758, 265010, 265011, 26448, 264763, 264683, 264288, 264685, 18108357, 29148659, 284690, 18108362, 264693, 18108370, 60431528, 18108374, 264634, 18108381, 56182323, 18108382, 18108388, 265526486, 87168518, 264482, 264487
2958	87440014 (5915, 5916)	87440014 (5915, 5916) Novel Protein sim. GBank gil4240257 dbj BAA74907.1 - (AB020691) KIAA0884 protein [Homo sapiens]			264595, 264596, 264681, 264369, 264629, 264631, 264567

2 (5929, 5930) NC gil rel	2965 80384762 (5929, 5930) Novel Protein sim. GBank gil4885447[ref]NP_005452.1[pKRML - Kreisler (mouse) markatel leucine zioner homolog		transcriptfactor	264259, 29331826, 264508, 264509, 264905, 264907, 264908, 264909, 264511, 265008,
				, 264910, , 264991, 264594, 33657402, 265011, 264760, 264762, 264764, 264763, 264686, 264626, 33657109, 264628, 264629, 33657109, 264631, 264631, 264634, 264633, 264636, 264637, 264638, 264636, 264636, 264638, 264688, 2646
<u>౽౿</u> ҈	91725248 (5931, 5932) Novel Protein sim. GBank gi[5262751 emb CAB45690.1 - (AJ243177) Xenopus RPA interacting protein alpha [Xenopus laevis]			60432289, 264682, 264448
<u> </u>	Novel Protein sim. GBank gij624225 (U19181) - Rabin3 Rattus norvegicus]		UNCLASSIFIED	264488, 264508, 264509, 264908, 264909, 264909, 264511, 264910, 264594, 264758, 85658542, 264762, 264764, 265021, 264556, 18108381, 264854
<u> </u>	Contains protein domain (PF00 gil4929715) Novel Protein sim. GBank CB-151881 CGI-123 Zinc finger, C3HC4 type (RING protein [Homo sapiens] Inger) Inger I	- (260	ерћ	264687, 52645156, 21906765, 52646365, 21906767, 18108398, 35696423, 22278996, 35696286, 22278997, 265020, 22278999, 265021, 265021, 265022, 264093, 264698, 264690, 262644150, 284259, 33857023, 5264690, 26444150, 284259, 33857023, 3264690, 264693, 29331824, 6671417, 29331825, 33109954, 52645182, 29331824, 6671417, 29331825, 33109954, 5264518, 29331826, 29331828, 265018, 26278900, 265019, 22278002, 26448, 66712502, 264569, 26448, 2646842, 22278996, 22278998, 26448, 66732049, 264269, 26331824, 29331826, 23331824, 263019, 26448, 264289, 265018, 265019, 26448, 264289, 265018, 265019, 26448, 264288, 264399, 265018, 265019, 26448, 264388, 264399, 265018, 265019, 26448, 264288, 264369, 265019, 26448, 264288, 264369, 265019, 26448, 264288, 264369, 265019, 26448, 264288, 264369, 2644150, 264418, 265011, 2618233, 264691, 33657199, 4618233, 264691, 33657199, 48108374, 56182333, 264891, 33657199, 4618233, 264891, 33657199, 4618233, 264891, 33657199, 46182323, 264891, 33657199, 4618233, 264891, 33657199, 4618233, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 46182333, 264891, 33657199, 461823333, 264891, 33657199, 46182333, 264891, 3665199, 46182333, 264891, 33657199, 46182333, 264891, 3665199, 46182333, 264891, 3665199, 46182333, 264891, 3665199, 46182333, 264891, 3665199, 46182333, 264891, 3665199, 46182333, 264891, 3665199, 46182333, 264891, 3665199, 4648, 264884, 264991, 3665199, 264891, 3665199, 46480, 264884, 2648988, 264898, 264898, 264898, 264898, 264898, 2648988, 2648988, 264898, 264898
S S	88088071 (5939, 5940) Novel Protein sim. GBank gi]3165407 (AC004755) - C [700807502_1 [140mo sapiens]	Contains protein domain (PF00046) - homeobox Homeobox domain		00110334, 01100310, 00432113, 22279000

264488, 56182575, 35696286, 56994075, 29331824, 29331826, 29146499, 264508, 264905, 264907, 264112, 264910, 21906754, 87168559, 265018, 265019, 18108351, 264689, 21906765, 21906767, 21906768, 265020, 60170615, 18108364, 264628, 264629, 18108384, 264636, 264558, 83373044, 18108384, 18108385, 87168518, 264564, 264567		IFIED 29331822, 264692, 33657349, 55811576, 264563	FIED	29331824, 29331827, 29331828, 264910, 85658542, 265011, 265018, 264448, 264288, 264769, 21906767, 265020, 264691, 264559, 83373044		IFIED 264509, 264288
7m1	kinase	UNCLASSIFIED	UNCLASSIFIED	ubiquitin	transport	UNCLASSIFIED
				Contains protein domain (PF00632) - ubiquitin HECT-domain (ubiquitin- transferase).		
94196930 (5941, 5942) Novel Protein sim. GBank gij728837 sp P39194 ALU7_HUMAN - II!! ALU SUBFAMILY SQ WARNING ENTRY !!!!	86625943 (5943, 5944) Novel Protein sim. GBank gi 728836 sp P39193 ALU6_HUMAN - III! ALU SUBFAMILY SP WARNING ENTRY III!	91215301 (5945, 5946) Novel Protein sim. GBank gil2746789 (AF040642) - No definition line found [Caenorhabditis elegans]	91673002 (5947, 5948) Novel Protein sim. GBank gil786117 (L41834) - nuclear protein [Ensis minor]	95325213 (5949, 5950) Novel Protein sin. GBank gil3890812 emb CAA19508 - Contains protein domain (AL023839) similar to HECT-domain (ubiquitin-transferase).; HECT-domain (ubiquitin-cDNA EST yk480410.5 comes from this gene [transferase].	87771202 (5951, 5952) Novel Protein sim. GBank gi[5679136 gb AAD46874.1 AF16093 - (AF160934) BcDNA.LD14189 [Drosophila melanogaster]) Novel Protein sim. GBank gij5262751 emb CAB45690.1 - (AJ243177) Xenopus RPA interacting protein alpha [Xenopus laevis]
94196930 (5941, 5942)	86625943 (5943, 5944)	91215301 (5945, 5946)	91673002 (5947, 5948)	95325213 (5949, 5950)	87771202 (5951, 5952)	91725254 (5953, 5954)
	2972	2973	2974	2975	2976	2977

		<u></u>	_		· · · · ·	П
2278995, 22278996, 22278997, 22278999, 264259, 60432289, 28331827, 29146499, 56182435, 265000, 265007, 265009, 265011, 265011, 265011, 265011, 265011, 265019, 264288, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 265022, 35686423, 264639, 265021, 265022, 35686423, 264539, 262021, 265022, 35686423, 264539, 262021, 265022, 35686423, 264539, 262021, 265022, 35686423, 264539, 262021, 265022, 262209002, 22279002	264488, 65274572, 56994075, 22278999, 264093, 29331822, 29331824, 264288, 55811957, 33657023, 33657109, 18108370, 55811576, 56182323, 60432113, 264482 265009, 21906767, 263981, 22279000	264629, 264564 22278995, 56994075, 22278996, 22278997, 22278998, 22278999, 264092, 29331824, 29331827, 29331828, 264092, 264591, 264592, 264594, 264595, 264596, 33657084, 264448, 21906765, 21906766, 21906767, 21906768, 21906769, 265022, 18108365, 33657182, 33657349, 35696423, 833373044, 22279000, 22279902	60424179, 52646842, 18108398, 22278997, 264093, 60432049, 264259, 29331822, 26032289, 33656970, 264905, 5264045, 265006, 60431735, 87168474, 265018, 265019, 18106351, 264448, 21906765, 21906768, 35695917, 33657023, 52645129, 18108370, 3569423, 83373044, 56526486, 60432113, 264404, 22279002		264910, 55812038, 56181562, 55811957, 264628, 55810764, 264632, 264635, 60432113	55811957, 264566 264369
UNCLASSIFIED	complement	UNCLASSIFIED		ATPase_associated	UNCLASSIFIED	UNCLASSIFIED
Contains protein domain (PF00480) - UNCLASSIFIED ROK family						
2978 87332059 (5955, 5956) Novel Protein sim. GBank gij746549 (U23522) - No definition line found [Caenorhabditis elegans]	91725256 (5957, 5958) Novel Protein sim. GBank gi[5262751 emb CAB45690.1 - (AJ243177) Xenopus RPA interacting protein alpha [Xenopus laevis] 86296600 (5959, 5960)	87376330 (5961, 5962) 95303675 (5963, 5964) Novel Protein sim. GBank 914929767[gb AAD34144.1 AF15190 - (AF151907) CGI-149 protein [Homo sapiens]	91725258 (5965, 5966) Novel Protein sim. GBank gi[\$262751 emb CAB45690.1 - (A/243177) Xenopus RPA interacting protein alpha [Xenopus laevis]	94136467 (5967, 5968) Novel Protein sim. GBank gi[2393734 (AC002542) - similar to C. elegans F11A10.5; 80% similarity to Z68297 (PID:g1130519) [Homo sapiens]	87099072 (5969, 5970) Novel Protein sim. GBank gi 103160 pir S22126 - finger protein unkempt - fruit fly (Drosophila melanogaster)	
87332059 (5955, 5956)		87376330 (5961, 5962) 95303675 (5963, 5964)			87099072 (5969, 5970)	86284861 (5971, 5972) 86455934 (5973, 5974)
978	2979 2980	2982	2983	2984	2985	2986 2987

2868 95357733 (5975, 5976) Nove Protein sim. Gaark gijde79028[pJAA027002.1] -	264488, 65274572, 22278995, 22278996, 22278996, 22278997, 224094, 264259, 60432049, 29331824, 29331826, 6043229, 35696052, 29331824, 29331826, 6042229, 35696052, 29331828, 264102, 264905, 264907, 264908, 66712502, 264828, 264909, 56182435, 265006, 265007, 265008, 60170831, 60432299, 264593, 60433356, 264757, 60433438, 21906754, 265010, 265011, 87168559, 265017, 265018, 264682, 256011, 265022, 263047, 265018, 264682, 256021, 265022, 263047, 265019, 265021, 265022, 263047, 265018, 264632, 265017, 265023, 265017, 265018, 264632, 264635, 264637, 283981, 264638, 264633, 263867, 283882, 264638, 264633, 264634, 26	83373044, 60432113, 22279000, 264563, 264564, 264565, 264566, 264567 22278996, 22278997, 264905, 264511, 60170831, 264593, 265019, 21906765,	22270390, 22270397, 2049U0, 264511, 6017031, 264593, 265019, 21906765, 21906765, 21906767, 21906768, 18108351, 264288, 264689, 265020, 264691, 33657023, 33657109, 264563	252-503 264-80. 265019, 264689, 18108385 264488, 29331822, 265017, 264761, 21906769, 65274791, 263981, 264565 22278994, 22278995, 56994075, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331827, 29331828, 265006, 265009, 264910, 33109954, 87168474, 87168559, 265019, 265019, 26448, 264288, 265021, 265022, 21906768, 21906766, 21906767, 21906768, 21906769, 265021, 26502, 33657023, 264693, 35695855, 83373044,	264905, 264907, 265019, 18108351, 264683 65274572, 35696286, 264259, 29331824, 35696052, 29146499, 264508, 264907, 265007, 265008, 60433438, 18108348, 265017, 274681, 264683, 264288, 264766, 264769, 264689, 35695917, 60170615, 33657023, 264692, 264634, 264555, 18108381, 18108362, 18108388, 264484
	UNCLASSIFIED	kinase	oncogene UNCLASSIFIED	UNCLASSIFIED UNCLASSIFIED UNCLASSIFIED transport	helicase
			Contains protein domain (PF00071) - Ras family		Contains protein domain (PF00270) - DEAD/DEAH box helicase
5976) 5980) 5980) 5990)	Novel Protein sim. GBank gij4679028 gb AAD27002.1 - (AF077207) HSPC021 [Homo sapiens]	Novel Protein sim. GBank gil113671[sp P23964 ALUF_HUMAN - !!!! ALU CLASS F WARNING FNTRY !!!	gil 1367 lisple 23964 JALUE HUMAN - IIII ALU CLASS F WARNING ENTRY III Novel Protein sim. CBank gil 2829836 isple 97348 IRHOD MOUSE - RHO-RELATED GTP-BINDING PROTEIN RHOD	Novel Protein sim. GBank gil2496549IsplQ50658IYU02_MYCTU - HYPOTHETICAL 29.7 KD PROTEIN CY339.02 Novel Protein sim. GBank qil2734081 (AE0001961) - similar	to oxysterol-binding proteins (Caenorhabditis elegans) Novel Protein sim. GBank gil2829912 (AC002291) - Similar ATP-dependent RNA Helicase (Arabidopsis thaliana)
2989 91225118 (5977, 2990 87330444 (5979, 2992 85425164 (5981, 2993 94325363 (5985, 2994 94136634 (5987, 2995 87591070 (5989, 12995 91013798 (5991, 12996 91013798 91013798 (5991, 12996 91013798 9101379 (5991, 12996 91013798 91013798 9101379 91013798 9101379 91	95357753 (5975, 5976)	91225118 (5977, 5978)	87330444 (5979, 5980) 94325381 (5981, 5982) 85425164 (5983, 5984)		

94) Novel Protein (AB023221) K	2997 87627440 (5993, 5994) Novel Protein sim. GBank gil4589652 db BAA76848.1 - (AB023221) KIAA1004 protein [Homo sapiens]	homeobox	264488, 56182575, 264259, 66714117, 29331826, 35696052, 264508, 264509,
			264907, 264908, 265006, 87168474, 265019, 264448, 264682, 264685, 264766, 21906764, 21906766, 21906766, 21906766, 21906766, 21906766, 21906768, 21906769, 27486261, 18108374, 35696423, 264634, 264635, 264637, 14108285, 8146848
2998 88095381 (5995, 5996) Novel Protein sim. G (AL034364) cDNA ES CDNA EST yk255b9. EMBL.M75923 come elegans]	Novel Protein sim. GBank gil3947589 emb CAA22252 - (AL034364) cDNA EST yk255b9.3 comes from this gene; cDNA EST yk255b9.5 comes from this gene; cDNA EST EMBL:M75923 comes from this gene [Caenorhabditis elegans]	UNCLASSIFIED	254636, 24337, 10106309, 87106319 52646365, 22778997, 264508, 264906. 18108351, 21906765, 21906767, 18108370, 18108374, 35696423, 264636, 264639
Novel Protein si gi 115408 sp P1 COLLAGEN 19	94847055 (5997, 5996) Novel Protein sim. GBank gi[115408 sp P18835 CC19_CAEEL - CUTICLE COLLAGEN 19	UNCLASSIFIED	56182575, 22278996, 28147620, 29331825, 29146498, 29146499, 264905, 66712502, 265006, 265009, 21906764, 85658642, 18108351, 29148627, 29148629, 60170615, 33657109, 27486262, 18108370, 18108374, 264565, 264558, 60170394,
Novel Protein yeast Sec6p. 8 to mammalian Method: conce norvegicus]	95099370 (5999, 6000) Novel Protein sim. GBank gi 1163174 (U32575) - similar to yeast Sec5p. Swiss-Prot Accession Number P32844; similar to mammalian B94, Swiss-Prot Accession Number Q03169; Method: conceptual translation supplied by author [Rattus norvegicus]	UNCLASSIFIED	264887, 22278997, 22278999, 264259, 26331822, 29331824, 35696052, 29146498, 264508, 264905, 264906, 264907, 264906, 264907, 264906, 265007, 264908, 264909, 264510, 265006, 265007, 265008, 264909, 264910, 33657402, 264757, 264690, 265017, 265000, 265017, 265008, 264761, 265008, 264761, 264762, 264681, 264682, 264761, 264769, 264681, 264769, 264769, 264681, 264769, 264681, 264692, 264681, 264769, 264691, 264692, 264691, 264692, 27486261, 264693, 264691, 264693, 264631, 264637, 264631, 264634, 264562, 264634, 264563, 264637, 264637, 264634, 264563, 264637, 264637, 264637, 264637, 264639, 264633, 264637, 264562, 264639, 264633, 264637, 264562, 264639, 264633, 264634, 264563, 264637, 264562, 264639, 264563, 264563, 264563, 264563, 264563, 264563, 264563, 264563, 264563, 264563, 264563, 264563, 264563, 264563, 264563, 264567, 26279002, 264563, 264486, 264567, 22279002, 264563, 264486, 264567, 22279002, 264563, 264486, 264567, 22279002, 264563, 264486, 264567, 22279002, 264563, 264486, 264567, 22279002, 264563, 264567, 26467,
Novel Protein sim. GE Putative gene. Gensc splicing.; coded for by (NID:g1678048), D31 (NID:g1733515), R59 (NID:g709111) [Homc	88078454 (6001, 6002) Novel Protein sim. GBank gi 2078470 (AC002073) - Putative gene. Genscan predictions confirmed by EST splicing.; coded for by human cDNAs AA122029 (NID:g1678048), D31562 (NID:g644442), AA158721 (NID:g1733515), R59640 (NID:g830335) and F13082 (NID:g709111) [Homo sapiens]	cathepsin	18108394, 52846842, 56182575, 29331824, 29331825, 29331825, 29331827, 264910, 33109954, 52644296, 265017, 265019, 264288, 265020, 265021, 52644150, 264692, 35695763, 56810764, 35696423, 56182323, 18108387, 264563, 264564
el Protein s	87718167 (6003, 6004) Novel Protein sim. GBank gi 3599478 (AF085185) - Myosin- A [Acanthamoeba castellanii]	UNCLASSIFIED	264488, 29331824, 29331825, 29331826, 29331827, 29331828, 264906, 264510, 265009, 21906754, 264682, 264688, 33657023, 264565

3003	85648079 (6005 6006)	3003 86648079 (6005 6006) Novel Protein sim GBank cit1754060 (1130303) collapse	Contractor at a state of a second		
		type XIII alpha-1 chain [Mus musculus]	Collagen triple helix repeat (20	conagen	264512, 264593, 264564, 264567, 264486
300	88066876 (6007 6008)	88066876 (6007 6008) Novel Protein sim CBank discoudabiles & 208001	copies)		
		(AB002342) KIAA0344 [Homo sapiens]	_		29331830, 21906769, 264691, 33657109, 263972 18108385
3005		87794843 (6009, 6010) Novel Protein sim. GBank	Contains protein domain (PF01360) - nxvgenase	OXVGENSEE	203312, 10100003
_		gil4680659lgb AAD27719.1 AF13294 - (AF132944) CGI-10			265011, 265019, 21906766, 21906767
		protein [Homo sapiens]	1		21906768, 265020, 33657023, 33657349
000	_				60170394, 22279002, 264567
2002	6/42224 (6011, 6012) Novel Protein Sim. Gl	Novel Protein sim. GBank gi 3930525 (AF064447) - sex-	Contains protein domain (PF00023) - MHC	МНС	264259, 29331822, 264512, 21906754,
		determination protein homolog Fem1a [Mus musculus]	Ank repeat		265018, 264687, 21906765, 264691, 264555.
3007	90936005 (6013, 6014) Novet Profein sim G	Novet Profein sim GBank oi/2565059 (1180738) - CACH13	Contains protein demais (DE00006)	74-14-14-14	264556, 264558, 18108385
		[Homo sapiens]	Zinc finger C2H2 type	transcriptiactor	525445U7, 52545156, 65274572, 264909,
					264689, 60170615, 18108374, 20281152
					264636, 52644332
900	80416249 (6015, 6016)				264905, 264593, 264766, 264636
5005	91213387 (6017, 6018)	91213387 (6017, 6018) Novel Protein sim. GBank gi 3127193 (AF062389) - kidney- Contains protein domain (PF00501) - [synthase	Contains protein domain (PF00501) -	synthase	52646842, 56182575, 22278995, 22278996,
		specific protein [Raftus norvegicus]	AMP-binding enzyme		264259, 29331825, 29331826, 29331827,
					29331828, 35696052, 264508, 264509,
					264907, 56182435, 264511, 265007, 264512,
					265008, 264757, 264758, 55812038, 264759,
					33109954, 21906754, 265010, 265011,
					264600, 265017, 265018, 265019, 264760,
					18108351, 264288, 264369, 21906764,
					21906765, 21906767, 55811957, 265020,
					265021, 264691, 18108368, 27486262,
_					20281149, 18108370, 55811576, 264637,
	`				264556, 264557, 18108381, 264558,
					56182323, 264559, 18108385, 18108388,
9,00					22279002, 264486
		83317217 (0019, 002U) Novel Protein Sim. GBank	Contains protein domain (PF01923) - UNCLASSIFIED	UNCLASSIFIED	264686, 264687, 21906767, 21906769,
		9149273701991AAD33084.1JAF06797 - (AF067972) DNA	Protein of unknown function		55811957, 22278995, 35695917, 22278996,
		cytosine metnyitransterase 3 alpha [Homo sapiens]			22278997, 265020, 265021, 60170615,
					264692, 33657023, 29331822, 264693,
					18108364, 29331824, 33657109, 60432289,
					29331827, 27486261, 29331828, 264508,
					264909, 55811576, 35695855, 265008,
					264556, 60433438, 83373044, 18108387,
					65274727, 60432113, 265017, 22279000,
2044	04292503 46034 60301				265019, 264564, 264682, 264764
	94353331 (0021, 0022)	Andreas (out 1, out 2) Individed Protein Sim. Grank	Contains protein domain (PF00153) - transport	transport	35696052, 56182435, 264758, 21906754,
		19/15/05/21/19/15/15/2011 1/AFT1/883 - (AFT1/8838) CITIN;	Mitochondrial carrier proteins		265018, 264760, 264762, 18108351, 264682,
		aduit-briser type ii cirruilinemia protein [Homo sapiens]			264448, 21906766, 65274620, 18108374,
3012	87753087 (6023 6024)				264482, 264564
	21 1 2000 1 (0023, 0024)			UNCLASSIFIED	263972

264488, 263994, 36696286, 22278997, 264259, 29331824, 60424269, 66714117, 35696052, 264906, 264906, 264907, 264909, 264909, 56182435, 264511, 264512, 264910, 264991, 364592, 264511, 264512, 264910, 264591, 264595, 264596, 56812038, 264591, 264596, 264596, 56812038, 264769, 265018, 264604, 264760, 264682, 264683, 264686, 264769, 264687, 21906767, 35695917, 265020, 33657023, 264629, 264693, 33657109, 264628, 264639, 264631, 264632, 264632, 264631, 264632, 264638, 264632, 264631, 264632, 264638, 264632, 264631, 264632, 264638, 264632, 264631, 264632, 264633, 264633, 264634, 264563, 2646643, 2646643, 2646643, 2646643, 2646643, 2646643, 2646643, 26466443, 26466444, 26464444, 2646644	264760	22278995, 22278996, 22278997, 264259, 29331824, 29331828, 264906, 265007, 265008, 264910, 265011, 265017, 265019, 264591, 33657109, 18108370, 35695855, 26456, 264564	52644507, 52846842, 56994075, 52645080, 29331822, 29331824, 35696052, 33656970, 52640495, 2654596, 33657084, 265017, 265019, 52644229, 21906767, 35695917, 52644150, 33657023, 33657109, 27486261, 3365763, 33695763, 3369556	264488, 18108397, 22278996, 35696286, 22278999, 264259, 29331822, 60432289, 264908, 29331830, 264909, 56182435, 265006, 265007, 265008, 265009, 264591, 6043336, 60433438, 52646317, 21906754, 55811389, 265010, 265011, 87186559, 265017, 265018, 265019, 264588, 264887, 21906765, 21906766, 21906767, 21906769, 265020, 265022, 65274620, 5245129, 33657109, 33657182, 18108370, 263972, 18108336, 2246432, 23357109, 33657182, 18108370, 263872, 264404, 60432113, 22279000, 264567
. Iranscriptfactor	ATPase_associated 264760		UNCLASSIFIED	
Contains protein domain (PF00400) - Irranscriptfactor WD domain, G-beta repeat				
3013 91238799 (6025, 6026) Novel Protein sim. GBank gij3702286 (AC005787) - R33374_1 [Homo sapiens]	79877263 (6027, 6028) Novel Protein sim. GBank gij3878374[emb]CAA93081] - (Z68879) Similarity to Yeast Chi12p protein (PIR Acc. No. S54453); cDNA EST EMBL:D27950 comes from this gene; cDNA EST EMBL:D27949 comes from this gene; cDNA EST EMBL:D33447 comes from this gene; cDNA EST EMBL:D33316 comes from this gene; cDNA EST		87759945 (6031, 6032) Novel Protein sim. GBank gil1168819ispiP41733 CC91_YEAST - CELL DIVISION CONTROL PROTEIN 91	95011154 (6033, 6034) Novel Protein sim. GBank gilq589658 dbj BAA76851.1 - (AB023224) KIAA1007 protein [Homo sapiens]
	79877263 (6027, 6028)			3017 95011154 (6033, 6034)

2040	14072904 46025				
0 00	3010 11073031 (6035, 6036)				264558
3019	94148231 (6037, 6038)	94148231 (6037, 6038) Novel Protein sim. GBank gij3219332 (AC004020) -		oncogene	264569, 52644507, 18108394, 65274572,
		מיויין שליים אינים שליים ביים ביים ביים ביים ביים ביים ביים			55182575, 22278994, 22278995, 55994075,
					22278998, 22278999, 264259, 29331822,
					29331824, 60432289, 29331827, 264908, 56183135, 365837, 365888, 50133338
					00102430, 2000U/, 2000U9, 0U432229, 064500, 604000FC, FEB40000, 040067F4
					204333,0U433330,33012U30,219U6/34, 87168474 265011 87168550 265017
					265018, 265019, 264681, 18108351, 264448.
					264682, 264683, 18108354, 264685, 264687.
					264689, 21906766, 21906768, 21906769,
					52644150, 264690, 264691, 33657023,
					264692, 264693, 33657109, 52645129,
					33657349, 264629, 65274791, 264634,
					52644332, 56182323, 18108385, 87168518,
000	100 00 0000 F100 F00 F00 F00 F00 F00 F00				22279000, 22279002, 264563
2020	94318251 (6039, 6040)	943 15231 (bu39, bu4u) Novel Protein sim. GBank gij3414809 (AF061529) - rjs [Mus Contains protein domain (PF00415) - ATPase_associated	Contains protein domain (PF00415)		264488, 263994, 35696286, 264259, 264508,
			codesisting (DCC1)		204909, 204909, 204900, 204907, 204900, 264000, 264640, 264040, 60474620, 264600
					2043US, 2043TU, 2043TU, 0017403S, 2040UU, 2646N3 26476N 264762 264683 264763
					204003, 204/00, 204/02, 204002, 204/03,
					264/64, 264288, 264369, 264/66, 264687,
					264688, 264769, 55811957, 35695917,
					33657023, 264628, 35696423, 35695855,
	_				264630, 264632, 264634, 264635, 264636,
					264637, 264556, 264557, 264638, 264639,
					83373044, 18108385, 264564, 264567,
	\neg				264486
3021		80478512 (6041, 6042) Novel Protein sim. GBank gij3880889 emb CAB09005 -			264769, 264629, 264482
		(Z95559) cDNA EST yk236d4.5 comes from this gene; cDNA EST EMBL C13465 comes from this gene; cDNA			
		EST 4432906 5 comes from this case: CDNA EST			
		CEMSH45R comes from this gene [Caenorhabditis elegans]			
3022	87718500 (6043 6044)			000000	***************************************
				UNCLASSIFIED	264259, 29331826, 29331828, 264288, 264566
3023			Contains protein domain (PF00614) - UNCLASSIFIED	UNCLASSIFIED	264488, 22278995, 35696286, 22278997,
		gil416592 sp[P32323 AGA1_YEAST - A-AGGLUTININ	Phospholipase D. Active site motif		29331826, 35696052, 264907, 29331830,
		ATTACHMENT SUBUNIT PRECURSOR			52644045, 56182435, 60432229, 264592,
					60433356, 60433438, 264689, 21906767,
			-		55811957, 35695917, 265021, 18108376,
	_				263978, 264635, 264558, 22279000
3024	85675305 (5047, 5048)			UNCLASSIFIED	60432049, 264760, 21906769, 55811957,
200					35695917, 264690, 264555, 264559
3023					264593, 55811576
		a weak suppressor of a mutant of the subunit AC40 of DNA dependant RNA polymerase I and III (Saccharomyces			
		cerevisiae]			

	, g , g	. . .	i s		
22278996, 22278997, 264490, 29331825, 264111, 265007, 60170831, 265010, 87168559, 265019, 21906765, 29148627, 263967, 20281149, 20281069, 263975, 2653977, 20281071, 56526486, 22279000	18108394, 22278996, 32578996, 35696286, 22278997, 22278998, 60432049, 264259, 29331822, 29331824, 6671417, 29331825, 60432289, 28331827, 29331828, 35696052, 264905, 264907, 263007, 265008, 264909, 264910, 23657402, 264596, 265009, 264910, 33657402, 264596, 264600, 265011, 87168559, 264680, 265018, 18108351, 264687, 264769, 264689, 21906765, 21906765, 21906767, 21906765, 21906766, 21906767, 21906761, 18108374, 55811576, 35696423, 265020, 266031, 264693, 32674791, 265020, 265021, 265022, 2664150, 264692, 32749024, 356574791, 264639, 264636, 264556, 264536, 264638, 22279002, 2227900	22278995, 22278996, 22278997, 22278999, 264259, 29331824, 29331827, 35696052. 264908, 265007, 265008, 265009, 60170831, 21908754, 265011, 87168559, 265018, 264765, 264689, 21906765, 21906768, 21906769, 29148629, 3569517, 265021, 265022, 33657109, 27486265, 264628, 264629, 18108374, 35696423, 35695855, 264638, 60170394, 22279000, 22279000, 22279000, 264482, 264584	22278997, 22278999, 29331827, 264905, 264509, 264509, 264510, 264511, 264512, 87168474, 265019, 18108351, 21906768, 264534, 264690, 264693, 263969, 18108370, 264588, 22279000, 22279002, 264482		22278995, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 35696052, 264908, 265018, 21906765, 21906766, 21906767, 21906768, 265021, 263974, 18108374, 264558, 56526486, 22279000, 22279002
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED		UNCLASSIFIED
	Contains protein domain (PF01529) - UNCLASSIFIED DHHC zinc finger domain	Contains protein domain (PF00400) - UNCLASSIFIED WD domain, G-beta repeat		Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	
3026 87643662 (6051, 6052) Novel Protein sim. GBank gi 3024052 sp P97924 KARI_RAT	9484563 (6053, 6054) Novel Protein sim. GBank gil4929647[gb AAD34084.1 AF15184 - (AF151847) CGI-89 protein [Homo sapiens]	94231997 (6055, 6056) Novel Protein sim. GBank gil3080521 emb CAA18650 - (AL022599) hypothetical protein [Schizosaccharomyces pombe]		87544928 (6059, 6060) Novel Protein sim. GBank gi 3757726 emb CAA18782 - (AL022727) dJ80119.1 (olfactory receptor-like protein (ns6M1-1)) [Homo sapiens]	91677953 (6061, 6062) Novel Protein sim. GBank gil4530587]gb AAD22105.1 - (AF132000) TADA1 protein [Homo sapiens]
87643662 (6051, 6052) r	94844563 (6053, 6054)				
3026	3027	3028	3029	3030	3031

3033 863	(100)	M. musculus MERS and other AHPC/TSA proteins Glycosyl transferases group 1	Contains protein domain (PF00034) - S	symmase	ZZZ/8996, 35696Z86, Z64Z59, Z93318Z4,	
			Glycosyl transferases group 1		29331828, 264907, 29331830, 264758,	
		[Caenorhabditis elegans]			33109954, 87168474, 87168559, 265019,	
					264288, 21906769, 265021, 264693,	
					35696423, 35695855, 264636, 56182323,	
	05308321 /6065 6066) News Bratein aim	Manual Drawing and Charle			83373044, 87158518	Т
	, (9000, 6000)	ODBOTA ODBOTA TIPACTE ABBO (ABBO)	Contains protein domain (PFUUUZZ) - struct	Struct	35696286, 264259, 29331826, 35696052,	
		grigos istración de contrata d	Acill		264508, 264905, 264906, 264907, 264908,	
		protein 5, years) nomong			264909, 265008, 264591, 21906/54, 265010,	
					265019, 264681, 264369, 264768, 21906764,	
					21906768, 35695917, 33657023, 264628,	
					35695855, 264632, 264635, 264639, 264482,	
3034 804	80415373 (6067, 6068)			INC. ACCIEIED	264563 264006 264007 264510 264603 265010	
	(2000)			ONCEASSIFIED	204900, 204901, 204310, 204392, 203010, 264762 264766 264637 264638 264486	
3035 912	20692 (6069, 6070)	91220692 (6069, 6070) Novel Protein sim. GBank gil3738207 emb CAA21262 -		UNCLASSIFIED	264636	T
		(AL031853) conserved ATP-GTP binding protein				
		[Schizosaccharomyces pombe]				
3036 917	718323 (6071, 6072)	91718323 (6071, 6072) Novel Protein sim. GBank	×	kinase	264907, 33657402, 265021	т
		017288371sofP391941ALU7 HUMAN - 1111 ALU SUBFAMILY				
		SQ WARNING ENTRY IIII				_
3037 953	107434 (6073, 6074)	95307434 (6073, 6074) Novel Protein sim GRank nil4406590inhlAAD200401.			265017	Т
		(AF131766) Similar to Ena-VASP like protein I Homo			7,000	
3038 954	21807 (6075, 6076)	otein sim GBank	Contains protein domain (BE00627) [INC] ASSIGIED		22278006 2278007 264250 264006	\neg
		14 14 14 14 14 14 14 14 14 14 14 14 14 1	- (1700 L.) Hallon majorid summing		22210990, 22210991, 204239, 2049U3,	_
		gi p35bu93 g0 AAD42865.1 AF15509 • (AF155099) NY-REN UBA domain	UBA domain		265007, 265009, 60433356, 21906754,	
		to antigen (nomo sapiens)			265018, 265019, 18108351, 264687,	
					21906765, 265020, 265021, 65274620,	_
`-					27486262, 264636, 56182323, 18108385,	
3030 873	97332757 6079\ No.:0 Draining	Note to the second of the second seco		T	222/9000	-1
	10700,1700)	Novel Frotein Sim. GBank gil475/128jembjCAB42094.1j -		UNCLASSIFIED	35696286, 29331828, 264109, 264110,	_
		(Autobor 17) the protein [Kaitus norvegicus]			264511, 265007, 21906754, 265011, 264681,	
					264683, 264687, 21906768, 264691,	
					18108370, 263972, 264629, 18108374,	
3040 000	133517 (6070 6000)	Alond Darking in Co. 11 - 12 control of the Control			263977, 35696423, 264564, 18108391	П
200	22211 (00/18, 0000)	adabasi (ours, outly) Novel Protein sim. Chank gil46642/8 emp CAB4324/.1			264692, 264558, 18108382, 18108385,	
2041 882	10367 (6001 6003)	(ALUSUU37) hypothetical protein [Homo sapiens]			264567	- 1
	12337 (0001, 0002)	oos (2337 (aug.), augz.) Novel Protein sim. GBank gij38/60/3jembjCAB04122.1j-		UNCLASSIFIED	56994075, 22278997, 22278998, 29331827,	
		(281505) similar to Zinc tinger, C3HC4 type (RING finger);			33656970, 33109954, 21906754, 87168559,	
		CUNA EST EMBL: U28025 comes from this gene; cDNA			264600, 264683, 21906765, 21906768,	
		EST EMBL:D28024 comes from this gene; cDNA EST			22279002	
		EMBL:D33210 comes from this gene; cDNA EST				
\neg		EMBL:D33441 comes from this				
3042 857	49402 (6083, 6084)	85749402 (6083, 6084) Novel Protein sim. GBank gi[790236 (U21156) -	5.	glycoprotein	264636	
		sarcolemmal associated protein-2 [Oryctolagus cuniculus]				\neg

3043	3043 87773026 (6085, 6086) Novel Protein sim. Gi (X83413) U88 Huma	Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6]		UNCLASSIFIED	35696286, 60424269, 35696052, 264508, 264905, 66712502, 56182435, 55811386, 52644296, 55811150, 35695917, 60170615, 33557109, 18108374, 264634, 60431850
3044		Novel Protein sim. GBank gild104922 (AF042276) - 0251 homolog [Pseudomonas putida]	Contains protein domain (PF01209) - glycoprotein ubiE/COQ5 methyltransferase family		22278996, 22278998, 22278999, 29331824, 56182435, 264511, 265007, 60170831, 60432229, 60433356, 33109954, 18108351, 264288, 35695917, 18108368, 18108370, 60170394
3045	94127598 (6089, 6090) Novel Protein sim. Gl (AB023232) KIAA101	Novel Protein sim. GBank gil4589680 dbj BAA76859.1 - (AB023232) KIAA1015 protein [Homo sapiens]	Contains protein domain (PF00096) - dna_rna_bind Zinc finger, C2H2 type		264488, 264259, 35696052, 264508, 264905, 264509, 264509, 264509, 264506, 264907, 265009, 264511, 265006, 264591, 265006, 264591, 264583, 33109954, 264604, 264764, 264683, 264288, 264766, 264768, 21906768, 25811957, 35695917, 27486262, 18108370, 264628, 18108374, 35695855, 264566
3046	3046 88098247 (6091, 6092)			UNCLASSIFIED	22278999, 29331822, 29331824, 29331825, 29331826, 29331826, 60432289, 29331827, 29331828, 264906, 52646317, 55811957, 60432113, 22279000, 22279002, 264482, 264564
3047				UNCLASSIFIED	264488, 22278996, 22278997, 22278999, 29331824, 29331825, 56182435, 264511, 265008, 265009, 265011, 265017, 264766, 21906768, 21906769, 35695917, 52644150, 33657349, 65274791, 35695855, 264555, 60432113, 22279000, 264566
3048	87629419 (6095, 6096)	87629419 (6095, 6096) Novel Protein sim. GBank gil4588034[gb AAD25962.1 AF09287 - (AF092878) zinc RING finger protein SAG [Homo sapiens]	Contains protein domain (PF00097) - UNCLASSIFIED Zinc finger, C3HC4 type (RING finger)		264102, 29148784
3049	86229955 (6097, 6098)	88229955 (6097, 6098) Novel Protein sim. GBank gij5454158 ref NP_006286.1 pVARS - valyI-IRNA synthetase 1	Contains protein domain (PF01406) - UNCLASSIFIED IRNA synthetases class I (C)		22278997, 29331826, 264907, 264758, 87168559, 265018, 264448, 21906766, 265020, 33657109, 35695855, 60432113, 22279000
3050		87643679 (6099, 6100) Novel Protein sim. GBank gil4589642 dbj BAA76843.1 - (AB023216) KIAA0999 protein [Homo sapiens]	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain		264259, 29331825, 264909, 265007, 264512, 265019, 264288, 21906766, 265020, 264693, 18108385, 56526486, 87168518, 22279002, 264566
3051					22278997, 264595, 265019, 264288, 264693, 87168518
3052	57108030 (6103, 6104)	57108030 (6103, 6104) Novel Protein sim. GBank gij117528 sp p14755 CRYL_RABIT - LAMBDA- CRYSTALLIN		dehydrogenase	264534

3053	3053 95350373 (6105, 6106) Novel Protein sim. GB	Novel Protein sim. GBank gij3947613 emb CAA19465.1 -	UNCLASSIFIED	65274572, 56181686, 22278995, 35696286, 22278998, 264759, 664759	
		gene; cDNA EST yk282d3.5 comes from this gene		265009, 60433438, 21906754, 265010,	
		[Caenorhabditis elegans]		87168559, 264603, 265018, 265019, 264763,	
				264764, 264288, 21906765, 21906766,	
				21906768, 21906769, 35695917, 181083/4,	
				35696423, 264638, 56182323, 22279000. 264563	
3054	86943510 (6107, 6108)	86843510 (6107, 6108) Novel Protein sim. GBank gil10762111bitil1550755 -	UNCLASSIFIED	35696286, 35696052, 29331830, 264908,	-
		hypothetical protein VSP-3 - Chlamydomonas reinhardtii		264909, 264512, 264910, 265017, 264604.	
				264766, 265020, 33657109, 264628,	
				35695855, 264636, 264564, 264566, 264486	
3055	95350537 (6109, 6110)	95350537 (6109, 6110) Novel Protein sim. GBank	transport	60424179, 65274572, 56182575, 35696286,	
,		gj 4680655 gb AAD27717.1 AF13294 - (AF132942) CGI-08		22278996, 22278999, 60432049, 264259,	
		protein [Homo sapiens]		60424269, 60432289, 35696052, 56182435,	_
				265006, 265009, 60170831, 60432229,	
				60431735, 60433356, 264594, 60433438.	
				21906754, 55811386, 265011, 87168559,	
				265019, 18108351, 264683, 264288, 264369,	
				264689, 21906768, 55811957, 35695917,	
				60170615, 33657023, 65274620, 33657109,	
				35695763, 60431528, 18108374, 55810764,	
				55811576, 35696423, 65274791, 264636,	
				60431850, 18108381, 56182323, 60170394,	
				18108385, 60432113, 264564, 264565,	
				264566	
3056	91661636 (6111, 6112)	91661636 (6111, 6112) Novel Protein sim. GBank	glycoprotein	264488, 264569, 18108394, 52646842,	
		gi 728837 sp P39194 ALU7_HUMAN - !!!! ALU SUBFAMILY		22278997, 22278998, 22278999, 264259.	
		SQ WARNING ENTRY !!!!		66714117, 29331826, 29331827, 35696052,	_
				264508, 264509, 264905, 264906, 264907,	
				264908, 264909, 265006, 264512, 265007,	
				265008, 265009, 264910, 33657402,	_
				55812038, 264596, 264758, 265010, 265011,	_
				265017, 265019, 264760, 18108351, 264762.	
			_	264763, 264764, 264288, 264766, 264687,	
				18108357, 264768, 264769, 264689,	
				21906765, 21906766, 21906767, 21906768,	
				21906769, 35695917, 265020, 265021,	
				264691, 264693, 33657109, 18108370,	_
				264628, 264629, 18108374, 55811576,	
				35696423, 35695855, 264630, 264631,	
				264632, 264634, 264635, 264636, 264537.	
				264638, 18108381, 83373044, 18108385,	
				22279000, 22279002, 264563, 264564,	
				264565, 264566, 264486, 264567	\neg

264508, 264905, 264907, 264908, 264909, 264510, 264512, 264910, 264592, 264594, 264767, 18108374, 264635, 264555, 26486 264639, 264563, 264565, 264486	264693	22278995, 22278996, 22278997, 22278998, 22278999, 264259, 29331824, 29331825, 29331827, 35696052, 29331828, 265607, 66433438, 265017, 265018, 265019, 264681, 264448, 264288, 264768, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 21906765, 2644150, 18108370, 264536, 18108385, 264563, 264567	264567	264112, 52644296, 21906768, 33657023, 263974, 18108385	264908, 265008, 18108351, 264566	18108359, 264558	52646365, 52646842, 65274572, 56182575, 56181686, 22278995, 22278996, 22278997, 22278996, 22278997, 22278999, 224259, 60432049, 29331824, 66774117, 264508, 264907, 264908, 25812038, 2564289, 265618, 264687, 264689, 21906768, 264691, 264689, 21906768, 264691, 264689, 21906768, 264691, 264636, 26182323, 264539, 18108370, 18108377, 55811576, 264693, 22279000, 2227900, 2227900, 2227900, 22279000, 22279000, 222790, 2227900, 2227900, 222790, 2227900, 222790, 22	22278996, 22278998, 264093, 264094, 264095, 29331824, 60424269, 66714117, 264100, 264907, 255007, 264591, 6043229, 264767, 21906765, 21906768, 264693, 20281069, 22279000, 22279002, 264482, 264566, 264567, 21906768, 264482, 264566, 264567
siruct	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	transferase	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED
		Contains protein domain (PF00097) - UNCLASSIFIED Zinc finger, C3HC4 type (RING finger)					Contains protein domain (PF00403) - UNCLASSIFIED Heavy-metal-associated domain	
3057 95412746 (6113, 6114) Novel Protein sim. GBank gi]3878119 emb CAA88860] - (249068) similar to GTP-binding protein; cDNA EST EMBL:M89111 comes from this gene; cDNA EST EMBL:D27709 comes from this gene; cDNA EST EMBL:D27708 comes from this gene; cDNA EST EMBL:D27708 comes from this gene; cDNA EST EMBL:D27708 comes from this gene; cDNA EST EMBL:D73788 comes from this gene; cDNA EST yk353.		87629425 (6117, 6118) Novel Protein sim. GBank gil4588034lgb AAD25962.1 AF09287 - (AF092878) zinc RING finger protein SAG [Homo sapiens]			Novel Protein sim. GBank gij4454690jgbjAAD20963j - (AF070657) glutathione S-transferase subunit 13 homolog [Homo sapiens]	80078023 (6125, 6126) Novel Protein sim. GBank gi[2246532 (U93872) - ORF 73, contains large complex repeat CR 73 (Kaposi's sarcomalassociated herpesvirus)	91241526 (6127, 6128) Novel Protein sim. GBank gi 4240315 dbj BAA74936.1 - (AB020720) KIAA0913 protein [Homo sapiens]	91639201 (6129, 6130) Novel Protein sim. GBank gi 5656743 gb AAD45960.1 AC00506 - (AC005067) Supported by Human EST H08032.1 (NID:g872854), mouse EST AA870042.1 (NID:g2965487), and genscan [Homo sapiens]
95412746 (6113, 6114)	79646226 (6115, 6116)	87629425 (6117, 6118)	79346691 (6119, 6120)	87740964 (6121, 6122)	87619465 (6123, 6124) Novel Protein sim. G (AF070657) glutathio [Homo sapiens]		91241526 (6127, 6128)	91639201 (6129, 6130)
257		3059	3060		3062	3063	3064	3065

18108397, 22278995, 56994075, 22278996, 264905, 66712502, 265006, 264512, 264910, 264758, 60174639, 264760, 18108351, 264764, 264683, 18108359, 264692, 18108364, 18108364, 18108379, 60170394, 264567	264488. 264489, 35696286, 22278996, 56994075, 264259, 29331822, 29331825, 35696052, 29331822, 29331825, 35696052, 29331822, 29331825, 2569005, 264509, 264508, 264508, 264908, 265019, 264511, 264512, 265008, 265019, 264511, 264512, 265008, 264594, 264511, 264512, 265008, 264594, 264591, 264592, 264593, 265010, 265011, 264512, 264603, 264604, 264096, 264766, 264766, 264764, 264766, 264762, 264687, 264769, 264689, 21906765, 21906767, 21906768, 264689, 21906765, 21906767, 21906768, 264689, 21906761, 264687, 264689, 264681, 38657023, 264639, 264639, 264634, 264634, 264635, 264636, 264634, 264536, 264637, 264638, 264638, 264636, 22279002, 262564, 264565, 264565, 264565, 264565, 264565, 264565, 264566, 264565, 264565, 264566, 264565, 264565, 264566, 264567, 264565, 264565, 264566, 264567, 264565, 264565, 264566, 264567, 264567, 264565, 264565, 264565, 264567, 264567, 264567, 264565, 264566, 264565, 264565, 264567, 264567, 264567, 264567, 264567, 264567, 264565, 264565, 264567, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467, 26467,	264112	22278996, 56994075, 22278998, 22278999, 264259, 264107, 264905, 29331830, 25644045, 264110, 60170831, 264592, 264594, 33657402, 21906754, 33109954, 8716859, 265017, 26448, 264764, 264683, 264766, 52644229, 21906765, 21906766, 21906768, 21906768, 21906768, 264557, 60170934, 56182323, 18108385, 87168518, 22279000, 264482
UNCLASSIFIED	struct		struct
	Contains protein domain (PF00787) - Struct PX domain		Contains protein domain (PF01926) - struct GTPase of unknown function
3086 91224437 (6131, 6132) Novel Protein sim. GBank gil4884268 emb CAB43245.1 - (AL050028) hypothetical protein [Homo sapiens]	95422551 (6133, 6134) Novel Protein sim. GBank gil4689258 gb AAD27832.1 AF12185 - (AF121859) sorting nexin 9 [Homo sapiens]		Novel Protein sim. GBank gij3878119jemblCAA88860j - (Z49068) similar to GTP-binding protein; cDNA EST EMBL.M89111 comes from this gene; cDNA EST EMBL:027709 comes from this gene; cDNA EST EMBL:027708 comes from this gene; cDNA EST EMBL:D7708 comes from this gene; cDNA EST EMBL:D73788 comes from this gene; cDNA EST yk353
91224437 (6131, 6132) (95422551 (6133, 6134)	85360651 (6135, 6136)	95412753 (6137, 6138)
3066	3067	3068	3069

264488, 22278994, 22278995, 262278996, 56994075, 22278997, 22278999, 264259. 29331822, 29147620, 29331824, 66714117, 29331826, 29147620, 29331824, 66714117, 29331826, 29146498, 29146499, 66712502, 265007, 264512, 264910, 60170831, 264592, 26458, 33109954, 21906754, 87168474, 265019, 18108351, 264448, 264683, 26428, 5264429, 265021, 60170615, 52644150, 264691, 33657023, 27486261, 27486264, 264628, 18108370, 18108377, 55811576, 35695855, 264634, 264628, 18108370, 18108377, 55811576, 35695855, 264634, 264658, 31373044, 18108385, 264568, 264404, 264563, 264566	264488, 65274572, 18108398, 22278996, 35666286, 22278997, 22278998, 224259, 22278999, 264259, 29331822, 66714117, 29331826, 264259, 29331828, 29146498, 29146499, 264907, 264908, 29331830, 264909, 264907, 264908, 29331830, 264909, 264907, 264908, 265009, 60170831, 60433229, 264952, 60433356, 33657402, 265010, 265017, 264681, 264288, 264685, 264766, 21906767, 21906769, 264697, 265021, 265022, 264692, 3365709, 264911, 264692, 3365709, 264911, 264692, 3365709, 264691, 356996423, 264635, 264636, 264630, 264630, 264631, 264635, 264636, 264630, 264630, 264631, 264635, 264565, 264560, 264500, 264563, 264563, 264565, 264566, 264567, 264563, 264565, 264566, 264567, 264567, 264567, 264566, 264567, 264567, 264567, 264566, 264567, 264567, 264567, 264566, 264567, 264567, 264567, 264567, 264566, 264567, 264567, 264566, 264567, 264670, 26467,	56182575, 29331822, 29331824, 29331825, 29146498, 264908, 52644045, 56182435, 265009, 60433438, 55812038, 18108351, 264683, 264369, 52644229, 52644150, 33657023, 264693, 33657109, 18108374, 55811576, 65274791, 264555, 56182323, 60432113, 264564
synthase	رۇز	transcriptfactor
	Contains protein domain (PF00085) - tgf	
3070 [94319173 (6139, 6140)] Novel Protein sim. GBank gi[3877788 emb CAB05527]- (Z83110) cDNA EST yk472b5.3 comes from this gene; cDNA EST yk474a7.3 comes from this gene; cDNA EST yk472b5.5 comes from this gene; cDNA EST yk468c10.3 comes from this gene; cDNA EST yk468c10.5 comes from this gene; cDNA EST yk468c10.5 comes from this gene; cDNA EST EM	Novel Protein sim. GBank gil4502425jref NP_001709.1 pBMP6 - bone morphogenetic protein 6 precursor	Novel Protein sim. GBank gil 1263289 (U47856) - fibroin-4 [Araneus diadematus]
70 94319173 (6139, 6140)	94325573 (6141, 6142)	3072 95115892 (6143, 6144) Novel Protein sim. [Araneus diademati

5/05		30/3 8b14/248 (b145, b145) Novel Protein sim. GBank gij134840 spjP22528 CORB_HUMAN - CORNIFIN B (SMALL PROLINE-RICH PROTEIN IB) (SPR-IB) (14.9 KD PANCORNULIN)		UNCLASSIFIED	264769
3074		88089351 (6147, 6148) Novel Protein sim. GBank gil3419847 (AC004982) - similar to yeast hypothetical protein ybk4; similar to P38164 (PID:g586461) [Homo sapiens]		UNCLASSIFIED	264488, 265019, 264488, 264288, 21906767, 264693, 18108378, 18108374, 264567
3075		88095752 (6149, 6150) Novel Protein sim. GBank gil4557349 ret NP_000456.1 pBARD - BRCA1 associated RING domain 1	Contains protein domain (PF00023) - homeobox Ank repeat	homeobox	264509, 264907, 264689, 264693, 56526486
3076				UNCLASSIFIED	18108398, 29331822, 29331827, 60432229, 265017, 264691, 264693
3077		88734277 (6153, 6154) Novel Protein sim. GBank gi 3023956 sp Q00808 HET1_PODAN - VEGETATIBLE INCOMPATIBILITY PROTEIN HET-E-1	Contains protein domain (PF00400) - kinase WD domain, G-beta repeat	kinase	65274572, 35696052, 264511, 60170831, 87168474, 264369, 35695917, 33657182, 27486264, 33657349, 35695763, 35695855, 264639
3078		88089355 (6155, 6156) Novel Protein sim. GBank gij3900850 (AC004994) - similar to KIAA0600; similar to d1026456 (PID:g3043724) [Homo sapiens]			22279002
3079		Novel Protein sim. GBank gij3875410jemb CAB02876 - (281052) Similarity to Yeast ABC1P protein (SW:ABC1_YEAST); cDNA EST yk229g8.3 comes from this gene; cDNA EST yk229g8.5 comes from this gene [Caenorhabditis elegans]		transport	29331824, 29331826, 264758, 55811386, 265017, 55811150, 52644229, 21906768, 265020, 265021, 264693, 18108376, 264631, 52644332, 22279002
0800		95298274 (6159, 6160) Novel Protein sim. GBank gi 5257221 gb AAD41265.1 - (AF117887) protein arginine methyltransferase [Mus musculus]		interferon	264489, 52644507, 22278996, 22278998, 264490, 264259, 29331824, 66714117, 29331825, 29331827, 29331828, 2946499, 264508, 264905, 264828, 25244045, 55182435, 265006, 264591, 264596, 21906754, 60174659, 264681, 264286, 264764, 264288, 264685, 264769, 264688, 264689, 21906757, 21906765, 21906767, 21906769, 25814150, 35695917, 265020, 60170615, 25844150, 264692, 33657109, 27486261, 35695763, 264628, 18108370, 62574791, 264568, 56182323, 60170334, 264688, 264688, 264689, 62744984
3081		88094864 (6161, 6162) Novel Protein sim. GBank gij728831 sp P39188 aLU1_HUMAN - !!!! ALU SUBFAMILY J WARNING ENTRY !!!!		UNCLASSIFIED	18108398, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265009, 264910, 264510, 264511, 265018, 264910, 264761, 264763, 264763, 264764, 18108354, 264685, 264763, 264629, 264630, 264631, 264632, 264634, 264635, 264638, 18108382, 18108385, 264563, 264565, 264586, 264688, 264663, 264663, 264663, 264663, 264663, 264663, 264663, 264664, 264665, 264668, 264668, 264668, 264668, 264668, 264668, 264668, 264668, 264668, 264668, 264668, 264668, 264668, 264668, 264668, 264668, 264688, 264668, 264689, 264688
3082	80310121 (6163, 6164)				264764, 55811957, 264555, 264564

60424179, 52645156, 65274572, 56182575, 56181686, 22278996, 32696286, 56994075, 22278996, 22278998, 22278999, 264259, 22378996, 22278999, 264259, 22331822, 56182181, 29331824, 23331825, 29331822, 29331822, 29331822, 29331822, 29331828, 3569605, 23365970, 264908, 265007, 265008, 60170831, 6043229, 60433356, 33657402, 56812038, 264758, 21906754, 33109954, 52646317, 55811386, 52644296, 87168474, 265011, 87168559, 265017, 265018, 265019, 264288, 264369, 18108357, 264768, 2906766, 21906762, 21906764, 21906766, 21906766, 21906766, 21906766, 21906767, 21906768, 2168288, 22789002, 22279002, 264563, 264482	35696286, 29331822, 35696052, 264508, 264509, 264909, 264909, 264909, 264510, 26510, 265011, 264683, 264683, 264663, 264659, 264653, 264639, 264639, 264639, 264639, 264689, 264689, 264685, 264639, 264482, 264639, 26488
UNCLASSIFIED	UNCLASSIFIED
	10
	Novel Protein sim. GBank gi]3873932[emb]CAB01859] - (Z79596) Simialrity to Bovine aspartyl beta hydroxylase (TR:G162694); cDNA EST EMBL:D27916 comes from this gene; cDNA EST EMBL:D27915 comes from this gene; cDNA EST EMBL:D64881 comes from this gene; cDNA EST EMBL:D68139 comes from this gene; cDNA
3090 95342371 (6179, 6180) Novel Protein sim. [Mus musculus]	3091 95317424 (6181, 6182) Novel Protein sim. (Z79596) Simialrity (TR:G162694); cDl gene; cDNA EST EMBL:1 CDNA EST EMBL:1 EMBL:1 EST EMBL:1

264488, 60424179, 18108396, 22278995, 56994075, 22278996, 35696286, 22278997, 22278998, 60432049, 284259, 29331822, 29331824, 29331824, 29331824, 29331826, 29331826, 29331826, 29331826, 29331826, 29331826, 29331826, 29331830, 264509, 264906, 264509, 264907, 264510, 265009, 264909, 264909, 264910, 265001, 264910, 265001, 264910, 265001, 264910, 265001, 264910, 265011, 264601, 264602, 265017, 264603, 264604, 265018, 264603, 264017, 26401, 264010, 264101, 264010, 264101, 264010, 264101, 264010, 264101, 264010, 264101, 264010, 264101, 264010, 264101, 264010, 264101, 264010, 264101, 26410	264259, 29331824, 35686052, 264905, 265006, 60432229, 60431735, 264684, 264369, 264288, 264766, 21906767, 35696423, 83373044, 18108385	18108398, 264259, 60432289, 29331827, 264511, 264763, 264288, 264767, 265022, 264691, 264693, 65274791, 56182323, 264564, 264565,
ribosomalprot	UNCLASSIFIED	- collagen
Contains protein G5 Ribosomal protein S5		Contains protein domain (PF01161) Phosphatidylethanolamine-binding protein
BOJESZ_HUMAN - 40S RIBOSOMAL LLREP3 PROTEIN)	Novel Protein sim. GBank gi 5002587 emb CAB44347.1 - (Y17454) LSFR1 protein [Homo sapiens]	Novet Protein sim. GBank gil400734[sp]P31044[PBP_RAT - Contains protein domain (PF01161) - collagen PHOSPHATIDYLETHANOLAMINE-BINDING PROTEIN (23 Phosphatidylethanolamine-binding KD MORPHINE-BINDING PROTEIN) (P23K)
3092 95314592 (6183, 6184) Novel Protein sim. (3093 94318457 (6185, 6186) Novel Protein sim. (Y17454) LSFR1 p.	94316675 (6187, 6188) Novel Prolein sim. PHOSPHATIDYLE KD MORPHINE-BI
3092	3093	3094

3095	94848162 (6189, 6190)		Contains protein domain (PF01454) - UNCLASSIFIED MAGE family	UNCLASSIFIED	18108397, 56182575, 22278995, 35596286, 56994075, 22278997, 22278999, 264259, 60432049, 66714117, 29331825, 60432289, 35696052, 33656970, 29146499, 264508, 264905, 264509, 264510, 264901, 264510, 264501, 264501, 265007, 265009, 260110, 265011, 8716859, 265003, 265010, 265011, 8716859, 265017, 265018, 264764, 264681, 264681, 264681, 264682, 264691, 265027, 2644150, 264691, 264691, 264691, 264691, 264691, 264691, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 265021, 264691,
3096	87756128 (6191, 6192) Novel Protein sim. (AB018293) KIAA0	Novel Protein sim. GBank gij3882221 dbj BAA3470.1 - (AB018293) KIAA0750 protein [Homo sapiens]	Contains protein domain (PF00307) - struct Calponin homology (CH) domain	struct	22278996, 22278996, 22278997, 22278999, 29331824, 29331825, 29331826, 29331827, 33658970, 284908, 265008, 264910, 33657402, 265011, 265017, 265018, 264369, 21906766, 21906767, 21906768, 35695917, 265020, 60170615, 264691, 264692, 264693, 264636, 248656, 12486562, 18108370, 60431528, 264634, 264636, 264639, 22279000, 264566
3097		88264895 (6193, 6194) Novel Protein sim. GBank gij4468288jembjCAB37981j- (AL022395) dJ273N12.1 (PUTATIVE protein based on EST matches) [Homo sapiens]	Contains protein domain (PF00646) - UNCLASSIFIED F-box domain.	UNCLASSIFIED	264486, 29331822, 29331825, 60432289, 29331826, 35696052, 29331828, 29331830, 264594, 55812038, 33109954, 33657084, 87168474, 87168559, 52644229, 21906765, 21906767, 18108376, 35696423, 52644332, 264638, 60432113, 22279002
3098					264634, 264637, 264565
3099	91243325 (6197, 6198) Novel Protein sim. (D12621) cytochror	- -		cyto450	264488, 35696286, 29331822, 29331824, 29331825, 29331827, 265007, 265008, 265010, 265011, 265018, 265019, 18108357, 21906766, 265020, 265022, 55811576, 56182323, 22279002, 264563
3100	87602421 (6199, 6200) Novel Protein sim. rich proteoglycan 2	Novel Protein sim. GBank gi 1083764 pir B48013 - proline- rich proteoglycan 2 precursor, parotid - rat		UNCLASSIFIED	29331825, 60432289, 35696052, 264910, 60432229, 264592, 264288, 264693, 263967, 264635
3101	79602134 (6201, 6202)			UNCLASSIFIED	264908, 264693, 264628, 264630, 264632

3102	91220892 (6203, 6204)	3102 91220892 (6203, 6204) Novel Protein sim. GBank gi[5305706]gb AAD41781.1 AF12853 - (AF128536) cytoplasmic phosphoprotein PACSIN2 [Homo sapiens]	Contains protein domain (PF00018) - struct SH3 domain		35696286, 22278996, 22278999, 29331827, 35696052, 264909, 264512, 265008, 60170831, 60433356, 33109954, 18108351, 264684, 264689, 21906767, 60170615.
					264692, 33657023, 264638, 22279000, 264482, 264564
3103	90938004 (6205, 6206)	90938004 (6205, 6206) Novel Protein sim. GBank gil464564 sp P35292 RB17_MOUSE - RAS-RELATED PROTEIN RAB-17		UNCLASSIFIED	35695917, 264565
3104		87340633 (6207, 6208) Novel Protein sim. GBank gi 5032207 ref NP_005696.1 pTSSC - tumor-suppressing STF cDNA 6		UNCLASSIFIED	264259, 264684, 264532, 33657182, 264558
3105	94148603 (6209, 6210)				22278997, 264259, 29331824, 35696052,
					29331828, 264508, 264509, 264905, 264906, 264907, 264908, 264511, 264910, 264591
					264594, 264758, 264760, 264681, 264762,
_					264764, 264288, 264766, 264768, 264687,
					264769, 21906766, 21906768, 35695917,
					33657023, 264692, 264693, 264628, 264629,
					35695855, 264630, 264631, 264632, 264634,
					264635, 264637, 264638, 264639, 83373044,
					264404, 22279002, 264563, 264565, 264566,
90,0	05004400 0040	The state of the s			264486, 264567
9	95351416 (5211, 5212)	93361416 (6211, 6212) Novel Protein Sim. GBank gi 1938574 (U97190) - B0025.2			22278996, 22278997, 22278998, 22278999,
		gene product (Caenornabditis elegans)			264092, 264093, 264094, 29331822, 264906,
					264907, 264908, 52644045, 56182435,
					264112, 265008, 265009, 55812038, 265017,
					265018, 264683, 264686, 264687, 264768,
					52644229, 21906765, 21906768, 21906769,
					55811957, 265020, 265022, 264690,
					0204410U, 20403Z, 204033, 1010037U,
					18108388, 22279000, 264563
3107	95343272 (6213, 6214)	95343272 (6213, 6214) Novel Protein sim. GBank gij3341441 emb CAA76851 -			22278995, 22278996, 35696286, 22278997.
		(Y17794) winged-helix transcription factor [Gallus gallus]			22278999, 264091, 264093, 264259,
					29331822, 29331825, 29331826, 60432289,
					29331827, 29331828, 33656970, 264105,
					264512, 265009, 60433356, 60433438,
					265011, 265017, 265018, 21906765,
					21906766, 21906767, 21906769, 265021,
					264691, 33657109, 27486261, 27486265,
					18108370, 263972, 18108374, 55811576,
3					18108385, 56526486, 264482, 264487
90 F2	87340635 (6215, 6216)	8/34/05/35 (6215, 6216) Novel Protein sim. GBank gij6032207[ref]NP_005696.1 pTSSC - tumor-suppressing		UNCLASSIFIED	56182435, 264288, 264690, 264564
		SIT CUINA D			

264490, 264908, 265007, 264910, 264593, 264683, 264684, 264687, 21906767, 21906768, 264693, 18108370, 264629, 18108374, 264638, 22279000	264488, 65274572, 22278995, 22278997, 60432049, 264259, 29331822, 29331824, 29331825, 29331824, 29331825, 26432289, 29331827, 29331828, 264506, 265007, 265008, 265009, 265011, 87168559, 264600, 265017, 265018, 265019, 18108351, 264288, 264509, 265010, 265019, 18108351, 264288, 265020, 60170615, 264693, 652749200, 18108370, 264693, 18108384, 22279000, 264563, 18108390	264488, 18108398, 66712502, 265017, 265018, 265019, 264448, 21906767, 265020, 33657023, 18108365, 18108368, 35696423, 52644332, 18108385, 18108388		18108397, 22278999, 264259, 29331824, 35680052, 264907, 264757, 60433438, 87168559, 264763, 264448, 18108354, 264288, 21906767, 21906769, 35695917, 264289, 264690, 264691, 264692, 264693, 18108365, 18108361, 18108381, 18108384, 18108385, 18108388, 87168518, 22279000, 22279002	<u>264909, 56182435, 264910, 21906754</u>	60424179, 56182575, 264259, 29331824, 60424269, 29331826, 66712502, 264510, 265007, 60431735, 60433356, 55812038, 55811386, 265019, 264288, 264689, 21906769, 264691, 33657023, 264693, 264595, 222739000, 222739000, 22273900	264905, 264758, 21906764, 264690	264510, 265011, 18108351, 264288, 264689, 264691, 18108368, 18108372, 263981, 264558, 264564
struct	UNCLASSIFIED	transcriptfactor	homeobox	lm7	hydrolase	transcriptfactor	UNCLASSIFIED	UNCLASSIFIED
Contains protein domain (PF00096) - struct Zinc finger, C2H2 type		Contains protein domain (PF00096) - Zinc finger, C2H2 type	Contains protein domain (PF00046) - homeobox Homeobox domain		Contains protein domain (PF00702) - hydrolase haloacid dehalogenase-like hydrolase			
3109 94318461 (6217, 6218) Novel Protein sim. GBank gij5002587 emb CAB44347 1 - (Y17454) LSFR1 protein [Homo sapiens]	95090716 (6219, 6220) Novel Protein sim. GBank gi 1076211 pir S50755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii	Novel Protein sim. GBank gij3282231 (U75454) - C2H2 type Contains protein domain (PF00096) - Iranscriptfactor zinc finger protein [Homo sapiens] Zinc finger, C2H2 type	88043639 (6223, 6224) Novel Protein sim. GBank gij3900848 (AC005023) - match to EST AA361117 (NID:g2013436) [Homo sapiens]	88207098 (6225, 6226) Novel Protein sim. GBank gil2459910 (AF005856) - anon2A5 [Drosophila yakuba]	79843167 (6227, 6228) Novel Protein sim. GBank gil4966270[gb]AAB52261.2] - (U97002) similar to acyl-CoA dehydrogenases and epoxide hydrolases; Pfam domain PF00441 (Acyl-CoA_dh), Score=57.4, E-value=1.7e-16, N=2; contains similarity to Pfam domain PF00702 (Hydrolase), Score=57.4, E-yalue=1e-13, N=1 [C	94117996 (6229, 6230) Novel Protein sim. GBank gi 5032225 ref NP_005676.1 pWBSC - Williams-Beuren syndrome chromosome region 11		
94318461 (6217, 6218)	95090716 (6219, 6220)	87754512 (6221, 6222) Novel Protein sim. G zinc finger protein [H			79843167 (6227, 6228)		3116 79642855 (6231, 6232)	87771288 (6233, 6234)
3109	3110	3111	3112	3113	3114	3115	3116	3117

52645156, 52646842, 65274572, 56182575, 22278995, 22278996, 35696286, 22278997, 22278998, 22278999, 26931824, 26714117, 29331826, 29331827, 35696052, 29331828, 264905, 29331827, 35696052, 29331828, 264905, 264908, 29331830, 52644045, 56182435, 264408, 264501, 265007, 265008, 265009, 264757, 52646317, 21906754, 33657084, 52644286, 87168474, 87168559, 265017, 265018, 264684, 264288, 264766, 2644186, 264682, 264691, 33657023, 264693, 33657023, 264693, 33657023, 264693, 3365703, 264628, 18108370, 60431528, 18108374, 35698423, 65274791, 60170394, 83373044, 87168518, 222780000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 222780000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 22278000, 222780	265006, 264288	264288, 264486 264288, 264486	5264507, 52645156, 52646365, 52846842, 22278994, 56994075, 22278996, 22278999, 264259, 29331824, 29331827, 35696052, 5264045, 265008, 52646317, 87168474, 87168599, 2966765, 52644150, 33657023, 31818374, 32627	264638	18108392, 29331822, 29331824, 29331825, 264905, 265007, 55812038, 265019, 18108351, 264682, 264288, 264766, 21906764, 21906765, 21906768, 21906769, 55811957, 18108365, 18108364, 18108384, 18108384, 22279000, 22372000	264905	56161686, 264259, 66714117, 60432289, 29331826, 29331827, 264907, 264908, 264828, 265009, 6043336, 33657402, 60433438, 264758, 18108351, 264288, 29148627, 29448629, 33657023, 33657109, 18108382, 56556486
J.61		UNCLASSIFIED		UNCI ASSIFIED	kinase	INCI ASSIFIED	dehydrogenase
Contains protein domain (PF00008) - tgf	Contains protein domain (PF00328) - Histidine acid phosphatase				Contains protein domain (PF00780) - CNH domain		Contains protein domain (PF00106) - dehydrogenase short chain dehydrogenase
Novel Protein sim. GBank gil3880563 emb CAB01444.1 - (Z78018) predicted using Genefinder; similar to serine/threonline kinase; cDNA EST yk353d10.5 comes from this gene [Caenorhabdftis elegans]		Novel Protein sim. GBank gi 5019819 gb AAD37863.1 AF14315 - (AF143152) putative NADH oxidoreductase complex I subunit [Caenorhabditis elegans]	94110735 (6241, 6242) Novel Protein sim. GBank gild501877 ref NP_001088.1 pACR - acrosin		Novel Protein sim. GBank gi[2439517 (AC002563) - putative Contains protein domain (PF00780) - kinase RHO/RAC effector protein; 95% similarity to P49205 (PID:g1345860) [Homo sapiens]		91216607 (6249, 6250) Novel Protein sim. GBank gil4980826[gb]AAD35412.1IAE00171 - (AE001714) oxidoreductase, short chain dehydrogenase/reductase family [Thermotoga maritima]
	85728796 (6237, 6238)	07344040 (0239, 0240)			88083003 (6245, 6246)	87786899 (6247, 6248)	91216607 (6249, 6250)
3118	3119	3120	3121	3122	3123	3124	3125

3126	3126 95337205 (6251, 6252)			LINCI ASSIFIED	22278999 26449N 264259 60432040
					29331822 60432289 29146408 52644045
		-			56182434 265000 60433438 265040
					30102433, 203008, 80433438, 203010,
_					87168559, 265017, 265018, 55811150,
					264763, 264683, 264369, 264685, 29148629,
					33657023, 264693, 33657109, 18108374,
					55811576, 18108385, 60432113, 22279002
312/	312/ 91539233 (5253, 5254) Novel Protein sim.				35696286, 22278996, 22278999, 29331826,
		(AL021687) putative protein [Arabidopsis thaliana]			264908, 60433438, 87168559, 264604,
					21906765, 21906769, 33657023, 33657349,
					264629, 18108374, 18108377, 22279000,
3		- 1			22279002
3128	3128 87674330 (6255, 6256) Novel Protein sim.	Novel Protein sim. GBank gij3885828 (AF090133) - lin-7-A Contains protein domain (PF00595) - misc_channel	Contains protein domain (PF00595) -	misc_channel	22278996, 264259, 52644045, 265008,
		[Kattus norvegicus]	PDZ domain (Also known as DHR or		21906754, 265017, 265018, 21906768,
			GLGF).		18108376, 18108387, 22279000, 22279002
3128	87722412 (6257, 6258)	3129 87735412 (5257, 5258) Novel Protein sim. GBank gij3135273 (AC003058) -	Contains protein domain (PF00400) - kinase	kinase	56182575, 264259, 29331825, 29331828,
		hypothetical protein [Arabidopsis thaliana]	WD domain, G-beta repeat		52644045, 56182435, 60433356, 264600,
					264682, 264763, 264764, 264369, 264288,
					264686, 55811957, 264692, 33657023,
3					33657109, 60432113, 264564, 264566
000	3130 149933900 (6259, 6260) Novel Protein sim.	Novel Protein sim. GBank gi 3329465 (AF064553) - NSD1 protein [Mus musculus]			264636
3131	95351469 (6261, 6262) Novel Protein sim.	Novel Protein sim. GBank gij1848277 (U86136) -	Contains protein domain (PF00400) - UNCLASSIFIED	UNCLASSIFIED	56182575, 264259, 29331824, 264907.
		telomerase-associated protein TP-1 [Homo sapiens]	WD domain, G-beta repeat		56182435, 264594, 60433438, 55812038,
					33109954, 21906754, 33657084, 87168474,
					264448, 264766, 21906769, 55811957,
					265020, 265021, 265022, 60170615,
					33657023, 33657109, 33657182, 27486261,
					33657349, 65274791, 60170394, 56182323,
					83373044, 87168518, 264564

52644507, 52646842, 52646365, 65274572, 56182575, 22278994, 22278995, 35696286, 56994075, 22278996, 22278996, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 22278999, 23331824, 29331825, 60432289, 29331826, 29331827, 29331825, 564870, 60170831, 60432229, 60433356, 33657402, 55812038, 52644045, 56182436, 25644296, 85658542, 87168559, 265017, 265018, 265019, 264488, 21906768, 21906768, 21906768, 21906768, 21906768, 21906768, 21906768, 23695917, 265021, 60170615, 52644150, 265021, 60170615, 52644150, 265021, 27486265, 33657349, 35695555, 181083374, 18108376, 55811576, 35695855, 18108385, 18108387, 56526486, 87168518, 60432113.	22278994, 22278998, 264905, 265006, 265007, 87168559, 264760, 21906767, 18108374, 22279000, 22279002, 264563	264595, 264369, 264685, 264628, 264566	22278996, 264095, 29331826, 33657402, 18108348, 263974	22278998, 264259, 264828, 265006, 265008, 60433438, 265019, 264764, 264288, 264769, 264689, 265020, 27486262, 263972, 65274791, 264557, 264558	22278995, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331825, 29331825, 29331825, 29331825, 265008, 21906754, 87188474, 265011, 87168559, 265017, 265018, 265011, 18108351, 264082, 264769, 21906765, 21906767, 21
ubiquitin	polymerase		struct	transport	UNGLASSIFIED
Contains protein domain (PF00789) - ubiquitin UBX domain				Contains protein domain (PF00153) - Mitochondrial carrier proteins	
3132 95415459 (6263, 6264) Novel Protein sim. GBank gil4680647[gb AAD27713.1 AF13293 - (AF132938) CGI-03 protein [Homo sapiens]	87379414 (6265, 6266) Novel Protein sim. GBank gil4507613 ref nP_003738.1 pTNKS - TANKYRASE	94649816 (6267, 6268) Novel Protein sim. GBank gil1729827lsp P54633 TALA_DICDI - FILOPODIN (TALIN HOMOLOG)	86389356 (6269, 6270) Novel Protein sim. GBank gij3093478 (AF012927) - fibrinogen-binding protein (Streptococcus equi)	94845839 (6271, 6272) Novel Protein sim. GBank gi 627101 pir S44092 - probable Contains protein domain (PF00153) - transport (carrier protein c2 - Caenorhabditis elegans Mitochondrial carrier proteins	88257947 (6273, 6274) Novel Protein sim. GBank gij3342730 (AC005331) - R31341_1 [Homo sapiens]
95415459 (6263, 6264,	3133 87379414 (6265, 6266)	3134 94649816 (6267, 6268)	3135 86389356 (6269, 6270)	3136 94845839 (6271, 6272)	3137 88257947 (6273, 6274

3138 94130186 (6275, 6276) Novel Protein sim. GBank gil4406759lgb AAD20070 - (AC006836) hypothetical protein [Arabidopsis thaliana]
87325503 (6277, 6278) Novel Protein sim. GBank gi 228938 prf 1814452C - Hyp-rich glycoprotein [Zea diploperennis]
91222692 (6279, 6280) Novel Protein sim, GBank dig32lemblCAA377731 -
icle
87323564 (6281, 6282) Novel Protein sim. GBank gi 3213227 (AF035209) - putative
95419028 (6283, 6284) Novel Protein sim. GBank gij2498197 sp Q95245 C561_PIG
93331473 (0283, 6285) Novel Protein sim. GBank gil5420387/emb[CAB46679.1]
(AJZ43459) proteophosphoglycan (Leishmania major)

4	95336329 (6287, 6288)	3144 95336329 (6287, 6288) Novel Protein sim. GBank gil4884468 emb CAB43322.1 - (AL050225) hypothetical protein [Homo sapiens]			264488, 18108396, 22278996, 35696286, 22278997, 22278999, 29331826, 29331827,
					35696052, 29331828, 264106, 265006, 265007, 265009, 33657402, 85658542,
					265011, 18108351, 264448, 264369, 21906765, 21906766, 21906767, 265020,
_					265021, 52644150, 27486261, 18108370,
					18108374, 35696423, 56182323, 83373044,
7					22279000, 22279002, 264567
3145	86611657 (6289, 6290)	86611657 (6289, 6290) Novel Protein sim. GBank gij3879709[emb]CAB03330] - (281118) Similarity to Human endosomal protein P162		UNCLASSIFIED	18108397, 29331824, 29146499, 20281100, 265006, 55812038, 265010, 21906766,
		(TR:Q15075); cDNA EST EMBL.Z14487 comes from this			29148627, 21906769, 29148784, 264692,
		gene; cDNA EST EMBL:Z14556 comes from this gene;			33657023, 33657109, 35695763, 263981,
		cDNA EST EMBL:D27011 comes from this gene; cDNA EST EMBL:D27015 comes from t			56182323, 87168518
3146		gen	Contains protein domain (PF00169) -	struct	264259, 29331826, 29331828, 29331830,
		inducible gene mig-2 - human	PH domain		264510, 264511, 265007, 265009, 264600.
					265017, 18108351, 264448, 264369,
					21906766, 265021, 264692, 33657109.
					18108374, 35696423, 35695855, 60432113,
					264564
3147	94848512 (6293, 6294) Novel Protein sim.	GBank gi[3874279]emb[CAB07315.1] -	Contains protein domain (PF00702) - UNCLASSIFIED	UNCLASSIFIED	56181686, 35696286, 60432049, 264259,
			haloacid dehalogenase-like		56182181, 29331825, 60432289, 35696052,
		/k315e12.5	hydrolase		56182435, 265008, 264910, 60431735,
_		comes from this gene [Caenorhabditis elegans]			60433356, 60433438, 265010, 264448,
					264288, 265022, 33657023, 33657109,
					60431528, 65274791, 264631, 56182323,
					264404, 22279002
3148	95362169 (6295, 6296) Novel Protein sim.	Novel Protein sim. GBank		UNCLASSIFIED	35696286, 35696052, 264511, 85658542,
		gi[5225322 gb[AAD40851.1[AF08310 - (AF083108) sirtuin			87168474, 264764, 35696423, 264555,
		type 3 [Homo sapiens]			264556, 264557, 264558, 83373044, 56526486, 60432113
3149	95308548 (6297, 6298) Novel Protein sim.	Novel Protein sim. GBank gil4200446 (AF102777) - FYVE	Contains protein domain (PF01363) - leph	eph	29331822, 35696052, 264109, 29148629,
	-	finger-containing phosphoinositide kinase [Mus musculus]	FYVE zinc finger	•	18108381
3150	87655472 (6299, 6300) Novel Protein sim.		Contains protein domain (PF00043) - transferase	transferase	264259, 29331822, 29331824, 29331825,
	-	(Y17850) ganglioside-induced differentiation associated	Glutathione S-transferases.		29331827, 52646317, 264686, 35695855,
┱		culus]			56182323, 264639
3151	87772355 (6301, 6302) Novel Protein sim.	GBank gij172591 (M63577) - SFP1	Contains protein domain (PF00096) - oncogene	oncogene	29331822, 265008
Т		Saccharomyces cerevisiae	Zinc finger, CZH2 type		
3152	85698108 (6303, 6304)			UNCLASSIFIED	21906754, 87168559, 264605, 21906768, FORMATED 27486264 35606423 22270000
7					32044 30, 21400204, 33030423, 2221 3000

264488, 52646365, 35696286, 22278996, 22278997, 22278999, 60432049, 264259, 29331826, 60432289, 33656970, 264268, 264908, 33657402, 264599, 60433438, 87168474, 8716859, 264601, 265019, 264448, 264828, 264764, 284288, 264369, 21906765, 21906766, 21906767, 21906768, 21906765, 21906766, 21906767, 265022, 60170615, 52644150, 264690, 264691, 33657023, 65274620, 33657109, 18108370, 35695855, 264638, 60170394, 87168518, 60432113, 22279000, 222799002		29331828, 264509, 264906, 264908, 264510, 264511, 264512, 33657402, 264681, 264683, 33657023, 18108370, 264634, 264639, 18108385, 264563, 264486	56182575, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 66714117, 29331825, 29331827, 35586052, 52644045, 265007, 265009, 60170831, 6043229, 60433356, 21906754, 33109954, 87168474, 265010, 265017, 265018, 265019, 18108351, 264448, 264228, 264689, 21906766, 21906768, 21906766, 21906766, 21906769, 25279002	18108396, 264259, 29331826, 35696052, 29146498, 87168559, 265017, 264448, 264289, 264691, 18108366, 52645129, 35696423, 52644332	264488, 263974	56182575, 22278999, 29331824, 264106, 60433356, 264758, 265011, 87168559, 26448, 18108354, 264768, 21906768, 265020, 264691, 264692, 33657109, 18108374, 35696423, 264555, 60170394, 22279000	18108351, 264555, 264556, 264557, 264568, 264559
struct	ATPase_associated	UNCLASSIFIED	transferase		kinase	UNCLASSIFIED	struct
Contains protein domain (PF00400) - struct WD domain, G-beta repeat			Contains protein domain (PF00652) - Iransferase Similarity to lectin domain of ricin beta-chain, 3 copies.		Contains protein domain (PF00023) - kinase Ank repeat		Contains protein domain (PF00023) - struct Ank repeat
3153 95317299 (6305, 6306) Novel Protein sim. GBank gil4895041[gb]AAD32705.1]AF14395 - (AF143957) coronin- WD domain, G-beta repeat [3 [Mus musculus]]	3) Novel Protein sim. GBank gil4680661lgblAAD27720.1 AF13294 - (AF132945) CGI-11 protein [Homo sapiens]	3155 87762394 (6309, 6310) Novel Protein sim. GBank gij728837 sp P39194 ALU7_HUMAN - !!!! ALU SUBFAMILY SQ WARNING ENTRY !!!!	gls630076lgblAAD45621.1 AC00601 - (AC006017) N-gils630076lgblAAD45621.1 AC00601 - (AC006017) N-acetylgalactosaminyltransferase; similar to Q10473 (PID:g1709559) [Homo sapiens]		b) Novel Protein sim. GBank gi[5306064 gb AAD41895.1 AF15677 - (AF156778) ASB-3 protein [Homo sapiens]	(AJ243800) WSC4 homologue [Kluyveromyces lactis]	0) Novel Protein sim. GBank gij3930525 (AF064447) - sex- determination protein homolog Fem1a [Mus musculus]
95317299 (6305, 6306)	87718573 (6307, 6308) Novel Protein sim. gil4680661lgblAAC protein [Homo sap	87762394 (6309, 6310,		88259577 (6313, 6314)	80034118 (6315, 6316) Novel Protein sim. gi 5306064 gb AAI protein [Homo sap	94124114 (6317, 6318) Novel Protein sim. (AJ243800) WSC4	80221068 (6319, 6320) Novel Protein sim- determination prot
3153	3154	3155	3156	3157	3158	3159	3160

_		_	_	_
[264488, 22278995, 22278997, 22278998,	264259, 29331822, 60432289, 29331828,	52644045, 265017, 265018, 264448, 264288.	21906764, 21906767, 265020, 18108374,	264636, 264566
88074111 (6321, 6322)				
3161				

Table 2

Tissue ID	Tissue Name	Tissue Information	Disease Association
20281069	192xN	Protein-protein Interactions	Any
20281071	192xN	Protein-protein Interactions	Any
20281149	192xN	Protein-protein Interactions	Any
20281152	192xN	Protein-protein Interactions	Any
264111	276xN	Protein-protein Interactions	Any
264112	276xN	Protein-protein Interactions	Any
263966	384xN	Protein-protein Interactions	Any
263967	384xN	Protein-protein Interactions	Any
264110	552xN	Protein-protein Interactions	Any
18108379	5PH 52.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy , Congenital Adrenal Hyperplasia,
18108381	5PH 52.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108383	5PH 52.3 (B's Lyphoma- Raji)	Lymphoma derived from B cells	Blood cancers, hematopoeisis, leukemia
18108368	5PH 52.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108384	5PH 52.6 (Brain- Thalmus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders,
			neuropsychiatric disorders
18108394	5PH 53.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy , Congenital Adrenal Hyperplasia,
18108355	5PH 53.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108359	5PH 53.3 (B's Lyphoma-Raji)	Lymphoma derived from B cells	Blood cancers, hematopoeisis, leukemia
18108361	5PH 53.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108362	5PH 53.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108366	5PH 53.6 (Brain- Thalmus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders, neuropsychiatric disorders
18108354	5PH 54.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy , Congenital Adrenal Hyperplasia.
18108392	5PH 54.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108348	5PH 54.3 (B's Lyphoma- Raji)	Lymphoma derived from B cells	Blood cancers, hematopoeisis, leukemia
18108382	5PH 54.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108395	5PH 54.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108365	5PH 54.6 (Brain- Thalmus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders,
			neuropsychiatric disorders
18108397	5PH 55.1 (Adrenai Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy , Congenital Adrenal Hyperplasia,
18108398	5PH 55.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108364	5PH 55.3 (B's Lyphoma- Raji)	Lymphoma derived from B cells	
18108388	5PH 55.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108358	5PH 55.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
20281099	5PH 56.2 (MG63)		
20281100	5PH 56.3 (UtSMC)		
264404	5PH.1 (Brain)	Whole Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalceimia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxiatelangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection

264510	5PH.10 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264511	5PH.11 (Placenta)	Placenta	Infertility, birth defects
264512	5PH.12 (Thyroid)	Thyroid	Hyperparathyroidism, Hypoparathyroidism
264555	5PH.13 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic
			thrombocytopenic purpura, autoimmume disease, allergies,
			immunodeficiencies, transplantation, Graft vesus host,
1		1	
264556	5PH.14 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic
			thrombocytopenic purpura, autoimmume disease, allergies,
	-		immunodeficiencies, transplantation, Graft vesus host,
264557	5PH.15 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic
			thrombocytopenic purpura, autoimmume disease, allergies,
			immunodeficiencies, transplantation, Graft vesus host,
		•	
264558	5PH.16 (Bone Marrow)	Вопе Магтом	Hemophilia, hypercoagulation, Idiopathic
İ			thrombocytopenic purpura, autoimmume disease, allergies,
			immunodeficiencies, transplantation, Graft vesus host,
		ŀ	, , , , , , , , , , , , , , , , , , , ,
264559	5PH.17 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic
			thrombocytopenic purpura, autoimmume disease, allergies,
			immunodeficiencies, transplantation, Graft vesus host.
		1	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
264569	5PH.19 (One Fetal tissue and	Mixed	
20 13 03	two cell lines)	1	
264687	5PH.19.1 (fetal thymus -	Fetal Thymus	Hemophilia, hypercoagulation, Idiopathic
204007	CRL7046)	Clair Triginius	thrombocytopenic purpura, immunodeficiencies
264688	5PH.19.2 (hematopoetic stem	Hamatopoeitic stem cells	Leukemia, osteoporosis, post-chemotherapeutic stem cell
204000	cells - CRL2043)	incinatopoenie stem cens	repopulation
264689	5PH.19.3 (osteogenic sarcoma	Osteogenie Sarcoma	Sarcomas, osteoporosis, osteopetrosis
204007	cell lines - HTB96)	Ostcogenie Sarconia	Salcomas, osicoporosis, osicopenosis
264690	5PH.19.4 (Fetal Liver)	Fetal liver	Von Hippel-Lindau (VHL) syndrome,
204090	31 11:19:4 (I clai Elvel)	l ctal livel	Cirrhosis, Transplantation
264691	5PH.19.5 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension,
204071	31 11.1 × .3 (11cm t)	licat	Congenital heart defects, Aortic stenosis, Atrial septal
			defect (ASD), Atrioventricular (A-V) canal defect, Ductus
			arteriosus, Pulmonary stenosis, Subaortic stenosis,
	1		Ventricular septal defect (VSD), valve diseases, Tuberous
	1		sclerosis, Scleroderma, Obesity, Transplantation
			Services, Generalist, General, Transplanation
264692	SPU 10.6 (Spines)	Calana	Hamanhilia Ulmana milaina 131tii
204092	5PH.19.6 (Spieen)	Spleen	Hemophilia, Hypercoagulation, Idiopathic
			thrombocytopenic purpura, Immunodeficiencies, Graft vesus host
264693	SPH 10.7 (Pipuitana)	District	
204093	5PH.19.7 (Pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's
			disease, Stroke, Tuberous sclerosis, hypercalceimia,
]		Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-
			telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
24462	I COLLA CO		
264482	5PH.2 (Brain)	Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's
]		disease, Stroke, Tuberous sclerosis, hypercalceimia,
	1		Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy, Lesch-Nyhan syndrome, Multiple
]		sclerosis, Ataxia-
	1		telangiectasia, Leukodystrophies, Behavioral disorders,
	,		Addiction, Anxiety, Pain, Neuroprotection

264600	5PH.21 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalceimia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxiatelangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264601	5PH.22 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmume disease, allergies, immunodeficiencies, transplantation, Graft vesus host,
264602	5PH.23 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264603	5PH.24 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264604	5PH.25 (Lymph Node)	Lymph Node	Lymphedema , Allergies
264605	5PH.26 (Placenta)	Placenta	Infertility, birth defects
264634	5PH.28 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation
264635	5PH.29 (Fetal Kidney)	Fetal Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalceimia, Lesch-Nyhan syndrome
264483	5PH.3 (Bone Marrow)	Bone marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmume disease, allergies, immunodeficiencies, transplantation, Graft vesus host,
264636	5PH.30 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264637	5PH.31 (P)ancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264638	5PH.32 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264639	5PH.33 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalceimia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxiatelangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264484	5PH.4 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmume disease, allergies, immunodeficiencies, transplantation, Graft vesus host,
264758	5PH.44.1 (Kidney)	Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalceimia, Lesch-Nyhan syndrome
2 64 760	5PH.44.2 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation

264762	5PH.44.3 (Heart)	Heart	I Continue to the continue to
204702	Si ii.44.3 (fieatt)	near	Cardiomyopathy, Atherosclerosis, Hypertension,
j		ĺ	Congenital heart defects, Aortic stenosis, Atrial septal
			defect (ASD), Atrioventricular (A-V) canal defect, Ductus
	1		arteriosus, Pulmonary stenosis, Subaortic stenosis,
			Ventricular septal defect (VSD), valve diseases, Tuberous
İ			sclerosis, Scleroderma, Obesity, Transplantation
264764	5PH.44.4 (Prostate)	Prostate	Prostate Cancer
264766	5PH.44.5 (Spleen)	Spleen	Hemophilia, Hypercoagulation, Idiopathic
	(thrombocytopenic purpura , Immunodeficiencies, Graft
264768	5PH.44.6 (pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome , Alzheimer's
1204700	Stri. 44.0 (pituliary)	r nunai y	
ì			disease, Stroke, Tuberous sclerosis, hypercalceimia,
			Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy, Lesch-Nyhan syndrome, Multiple
			sclerosis,Ataxia-
1	1		telangiectasia, Leukodystrophies, Behavioral disorders,
264769	(SDU 44.7 (16)		Addiction, Anxiety, Pain, Neuroprotection
264905	5PH.44.7 (Uterus)	Uterus	Infertility, birth defects
204903	5PH.48.1 (Burkitt's Lymphoma- Raji)	Burkitt's Lymphoma	Lymphoma, blood cancers
264906	5PH.48.2 (Thalamus- Brain)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's
			disease, Stroke, Tuberous sclerosis, hypercalceimia,
			Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy, Lesch-Nyhan syndrome, Multiple
			sclerosis, Ataxia-
			telangiectasia, Leukodystrophies, Behavioral disorders,
			Addiction, Anxiety, Pain, Neuroprotection
264007	(DI 40.2 (4.1. 1.0)		
264907	5PH.48.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy , Congenital Adrenal Hyperplasia,
264908	5PH.48.4 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
264909	5PH.48.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
264910	5PH.48.6 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
265006	5PH.50.1 (B's lymphoma)	Burkitt's Lymphoma	Lymphoma, blood cancers
265007	5PH.50.2 (thalamus)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's
			disease, Stroke, Tuberous sclerosis, hypercalceimia,
			Parkinson's disease, Huntington's disease, Cerebral palsy,
		1	Epilepsy, Lesch-Nyhan syndrome, Multiple
			sclerosis, Ataxia-
			telangiectasia, Leukodystrophies, Behavioral disorders,
			Addiction, Anxiety, Pain, Neuroprotection
265008	5PH.50.3 (adrenal gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy , Congenital Adrenal Hyperplasia,
265009	5PH.50.4 (fetal lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
203009		Salivary Gland	Dry mouth, infection
265010	5PH.50.5 (salivary gland)	<u> </u>	
	5PH.50.5 (salivary gland) 5PH.50.6 (mammary gland)	Mammary Gland	Lactation disorders, breast cancer
265010 265011 18108385			Lactation disorders, breast cancer Breast Cancer
265010 265011	5PH.50.6 (mammary gland)	Mammary Gland	
265010 265011 18108385	5PH.50.6 (mammary gland) 5PH.51.1 (MCF-7) 5PH.51.2 (CCRF-CEM) 5PH.51.3 (K-562)	Mammary Gland Breast Cancer	Breast Cancer
265010 265011 18108385 18108370	5PH.50.6 (mammary gland) 5PH.51.1 (MCF-7) 5PH.51.2 (CCRF-CEM)	Mammary Gland Breast Cancer Cancer Cell line	Breast Cancer Cancer
265010 265011 18108385 18108370 18108374 18108351	5PH.50.6 (mammary gland) 5PH.51.1 (MCF-7) 5PH.51.2 (CCRF-CEM) 5PH.51.3 (K-562)	Mammary Gland Breast Cancer Cancer Cell line Cancer Cell line	Breast Cancer Cancer Cancer
265010 265011 18108385 18108370 18108374 18108351 18108372	5PH.50.6 (mammary gland) 5PH.51.1 (MCF-7) 5PH.51.2 (CCRF-CEM) 5PH.51.3 (K-562) 5PH.51.4 (OVCAR-3)	Mammary Gland Breast Cancer Cancer Cell line Cancer Cell line Ovarian cancer	Breast Cancer Cancer Cancer Ovarian cancer Cancer
265010 265011 18108385 18108370 18108374	5PH.50.6 (mammary gland) 5PH.51.1 (MCF-7) 5PH.51.2 (CCRF-CEM) 5PH.51.3 (K-562) 5PH.51.4 (OVCAR-3) 5PH.51.5 (HL-60)	Mammary Gland Breast Cancer Cancer Cell line Cancer Cell line Ovarian cancer Cancer Cell line	Breast Cancer Cancer Cancer Ovarian cancer Cancer Hemophilia, hypercoagulation.ldiopathic
265010 265011 18108385 18108370 18108374 18108351 18108372	5PH.50.6 (mammary gland) 5PH.51.1 (MCF-7) 5PH.51.2 (CCRF-CEM) 5PH.51.3 (K-562) 5PH.51.4 (OVCAR-3) 5PH.51.5 (HL-60)	Mammary Gland Breast Cancer Cancer Cell line Cancer Cell line Ovarian cancer Cancer Cell line	Breast Cancer Cancer Cancer Ovarian cancer Cancer

264508	5PH.8 (Fetal Brain)		Von Hippel-Lindau (VHL) syndrome, Alzheimer's
	, , , , , , , , , , , , , , , , , , , ,		disease, Stroke, Tuberous sclerosis, hypercalceimia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple
		İ	sclerosis, Ataxia- telangiectasia, Leukodystrophies, Behavioral disorders,
Ì			Addiction, Anxiety, Pain, Neuroprotection
264509	5PH.9 (Lymph Node)	Lymph Node	Lymphedema, Allergies
20798451	5RH 56.3(UtSMC)		
264487	5RH.I (Brain)	Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalceimia. Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264534	5RH.11 (Bone marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmume disease, allergies, immunodeficiencies, transplantation, Graft vesus host,
264535	5RH.12 (Bone marrow)	Вопе Магтоw	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmume disease, allergies, immunodeficiencies, transplantation, Graft vesus host,
264563	5RH.19 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalceimia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxiatelangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264488	5RH.2 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation.ldiopathic thrombocytopenic purpura, autoimmume disease, allergies, immunodeficiencies, transplantation, Graft vesus host,
264564	5RH.20 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264565	5RH.21 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264566	5RH.22 (Placenta)	Placenta	Infertility, birth defects
264567	5RH.23 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264591	5RH.25 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke. Tuberous sclerosis, hypercalceimia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxiatelangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264592	5RH 26 (Bone Marrow)	Вопе Магтоw	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmume disease, allergies, immunodeficiencies, transplantation, Graft vesus host,
264593	5RH.27(thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264594	5RH.28 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264595	5RH.29 (Lymph Node)	Lymph Node	Lymphedema , Allergies
264489	5RH.3 (Bone Marrow)	Вопе Магтоw	Hemophilia. hypercoagulation, idiopathic thrombocytopenic purpura, autoimmume disease, allergies, inununodeficiencies, transplantation, Graft vesus host,

264596	5RH.30 (Placenta)	Placenta	Infertility, birth defects
264628	5RH.33 (fetal Kidney)	Fetal kidney	Diabetes, Autoimmune disease, Renal artery stenosis,
			Interstitial nephritis, Glomerulonephritis, Polycystic
	l		kidney disease, Systemic lupus erythematosus, Renai
			tubular acidosis, IgA nephropathy, Hypercalceimia, Lesch-
			Nyhan syndrome
264629	5RH.34 (lymph Node)	Lymph Node	Lymphedema, Allergies
264630	5RH.35 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264631	5RH.36 (thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264632	5RH.37 (Fetal Brain)	Fetal Brain	Von Hippel-Lindau (VHL) syndrome , Alzheimer's
	,		disease, Stroke, Tuberous sclerosis, hypercalceimia,
			Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy, Lesch-Nyhan syndrome, Multiple
			sclerosis, Ataxia-
			telangiectasia, Leukodystrophies, Behavioral disorders,
		}	Addiction, Anxiety, Pain, Neuroprotection
24400	CONTRACTOR OF THE CONTRACTOR O		. 1
264490	5RH.4 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic
			thrombocytopenic purpura, autoimmume disease, allergies,
			immunodeficiencies,transplantation, Graft vesus host,
264681	5RH.43.1 (fetal thymus -	Fetal Thymus	Hemophilia, hypercoagulation, Idiopathic
204001	CRL7046)	l car mymus	thrombocytopenic purpura, immunodeficiencies
264682	5RH.43.2 (hernatopoetic stem	Hematopoeitic stem cells	Leukernia, osteoporosis, post-chemotherapeutic stern cell
204002	cells - CRL2043)	l stem cens	repopulation
264683	5RH.43.3 (osteogenic sarcoma	Osteogenic Sarcoma	Sarcomas, osteoporosis, osteopetrosis
	cell lines - HTB96)		
264684	5RH.43.4 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome,
			Cirrhosis, Transplantation
264685	5RH.43.6 (Spleen)	Spleen	Hemophilia, Hypercoagulation, Idiopathic
			thrombocytopenic purpura, Immunodeficiencies, Graft
			vesus host
264686	5RH.43.7 (pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's
	i		disease, Stroke, Tuberous sclerosis, hypercalceimia,
			Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy, Lesch-Nyhan syndrome, Multiple
	1		sclerosis,Ataxia-
			telangiectasia, Leukodystrophies, Behavioral disorders,
			Addiction, Anxiety, Pain, Neuroprotection
264757	5RH.44.1 (Kidney)	Kidney	Diabetes, Autoimmune disease, Renal artery stenosis,
			Interstitial nephritis, Glomerulonephritis, Polycystic
			kidney disease, Systemic lupus erythematosus, Renal
			tubular acidosis, IgA nephropathy, Hypercalceimia, Lesch-
	ļ		Nyhan syndrome
264759	5RH.44.2 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome,
			Cirrhosis, Transplantation
264761	5RH.44.3 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension,
			Congenital heart defects, Aortic stenosis, Atrial septal
			defect (ASD), Atrioventricular (A-V) canal defect, Ductus
	1		arteriosus, Pulmonary stenosis, Subaortic stenosis,
			Ventricular septal defect (VSD), valve diseases, Tuberous
			sclerosis, Scleroderma, Obesity, Transplantation
	CRIT 44 A CR		
264763	5RH.44.4 (Prostate)	Prostate	Prostate Cancer
264765	5RH.44.5 (Spleen)	Spicen	Hemophilia, Hypercoagulation, Idiopathic
			thrombocytopenic purpura, Immunodeficiencies, Graft
	1		· vesus host

264767	5RH.44.6 (Pituitary)	Interior	
204707	SKII.44.0 (Fituliary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's
			disease, Stroke, Tuberous sclerosis, hypercalceimia,
l			Parkinson's disease, Huntington's disease, Cerebral palsy,
l			Epilepsy,Lesch-Nyhan syndrome, Multiple
			sclerosis, Ataxia-
1			telangiectasia, Leukodystrophies, Behavioral disorders,
			Addiction, Anxiety, Pain, Neuroprotection
264828	5RH.46.1 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264887	5RH.47.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome,
			Cirrhosis, Transplantation
18108377	5RH.50.1 (B's lymphoma)	Burkitt's Lymphoma	Lymphoma, blood cancers
18108380	5RH.50.2 (thalamus)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's
			disease, Stroke, Tuberous sclerosis, hypercalceimia,
	1		Parkinson's disease, Huntington's disease, Cerebral palsy,
	 		Epilepsy, Lesch-Nyhan syndrome, Multiple
			sclerosis, Ataxia-
			telangiectasia, Leukodystrophies, Behavioral disorders,
			Addiction, Anxiety, Pain, Neuroprotection
18108396	5RH.50.3 (adrenal gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108391	5RH.50.4 (fetal lung)	Fetal Lung	Airway diseases, infection
18108357	5RH.50.5 (salivary gland)	Salivary Gland	Dry mouth, infection
18108390	5RH.50.6 (mammary gland)	Mammary Gland	Lactation disorders, breast cancer
264532	5RH.9 (Bone Marrow)	Вопе Магтом	Hemophilia, hypercoagulation, Idiopathic
	(Some Mariow)	Bone Mariow	thrombocytopenic purpura, autoimmume disease, allergies,
			immunodeficiencies, transplantation, Graft vesus host,
			minumodenciencies, transplantation, Graft vesus nost,
263974	736xN		
263976	736xN		
263981	736xN		
20281166	96xN		
20281169	96xN		
20281171	96xN		
263994	cDNA-ORF Selection		
264080	Mx96		-
21906754	NQH 6.1 (HH729)	† 	
22278996	NQH 6.10 (PrEC)	Endothelial ceils	heart disease, cancer
22278997	NQH 6.11 (CAEC)	Endothelial cells	heart disease, cancer
22278998	NQH 6.12 (CSC)	Cancer Cell line	Cancer
22278999	NQH 6.13 (NHNPC)	Cancer Cell line	Cancer
22279000	NQH 6.14 (NHMC-RM)	Cancer Cell line	Cancer
22279002	NQH 6.15 (Hypothalmus)	Hypothalamus	Von Hippel-Lindau (VHL) syndrome , Alzheimer's
	(11) politically	riypotitularitas	disease, Stroke, Tuberous sclerosis, hypercalceimia,
			Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy, Lesch-Nyhan syndrome, Multiple
		İ	sclerosis, Ataxia-
			telangiectasia, Leukodystrophies, Behavioral disorders,
2100655	Wolf of the second		Addiction, Anxiety, Pain, Neuroprotection, Obesity
21906764	NQH 6.2 (In Dated Platelets)	Platelets	Clotting diseases, stroke
21906765	NQH 6.3 (HuVec)	Endothelial cells	heart disease, cancer
87168474	NQH 6.3 (Sized-HUVEC)	Endothelial cells	heart disease, cancer
21906766	NQH 6.4 (UtMVEC- myo)	Cancer Cell line	Cancer
21906767	NQH 6.5 (NHEM-neo)	Cancer Cell line	Cancer
21906768	NQH 6.6 (NHEK)	Cancer Cell line	Cancer
_		<u> </u>	
21906769	NQH 6.7 (ByCAEC)	Endothelial cells	heart disease, cancer

22278995	NQH 6.9 (PrSC)	Cancer Cell line	Cancer
27486261	NQH 7.1 (Jurkat E6-	Cancer Cell line	Cancer
	untreated)		
27486262	NQH 7.2 (TF1-untreated)	Cancer Cell line	Cancer
27486264	NQH 7.3 (U87-untreated)	Cancer Cell line	Cancer
27486265	NQH 7.4 (THP1-untreated)	Cancer Cell line	Cancer
29331822	NQH 8.1 (Brain- amygdala)		Von Hippel-Lindau (VHL) syndrome, Alzheimer's
			disease, Stroke, Tuberous sclerosis, hypercalceimia,
			Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy, Lesch-Nyhan syndrome, Multiple
			sclerosis,Ataxia-
			telangiectasia, Leukodystrophies, Behavioral disorders,
l .			Addiction, Anxiety, Pain, Neuroprotection
29331824	NQH 8.2 (Brain-		Von Hippel-Lindau (VHL) syndrome, Alzheimer's
	hippocampus)		disease, Stroke, Tuberous sclerosis, hypercalceimia,
ł			Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy, Lesch-Nyhan syndrome, Multiple
l			sclerosis,Ataxia-
	1		telangiectasia, Leukodystrophies, Behavioral disorders,
	ľ		Addiction, Anxiety, Pain, Neuroprotection
29331825	NQH 8.3 (Brain-substantia		Von Hippel-Lindau (VHL) syndrome, Alzheimer's
l	nigra)		disease, Stroke, Tuberous sclerosis, hypercalceimia,
	1		Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy,Lesch-Nyhan syndrome, Multiple
			sclerosis,Ataxia-
1			telangiectasia, Leukodystrophies, Behavioral disorders,
Í			Addiction, Anxiety, Pain, Neuroprotection
29331826	NOH 8.4 (small intestine)	C11 :	
- 2001666	14Qri 8.4 (Smail miestine)	Small intestine	digestive diseases, obesity, diabetes
29331827	NQH 8.5 (Spinal cord)	Spinal chord	digestive diseases, obesity, diabetes paralysis, neurodegenerative disorders
			1.3
29331827	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea)	Spinal chord	paralysis, neurodegenerative disorders
29331827 29331828	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG-	Spinal chord Stomach	paralysis, neurodegenerative disorders Stomach cancer
29331827 29331828 29331830 87168518	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool)	Spinal chord Stomach	paralysis, neurodegenerative disorders Stomach cancer
29331827 29331828 29331830	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2	Spinal chord Stomach	paralysis, neurodegenerative disorders Stomach cancer
29331827 29331828 29331830 87168518	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated)	Spinal chord Stomach	paralysis, neurodegenerative disorders Stomach cancer
29331827 29331828 29331830 87168518	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2	Spinal chord Stomach	paralysis, neurodegenerative disorders Stomach cancer
29331827 29331828 29331830 87168518 87168559 35695763	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated)	Spinal chord Stomach Trachea Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer
29331827 29331828 29331830 87168518 87168559	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment	Spinal chord Stomach Trachea	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection
29331827 29331828 29331830 87168518 87168559 35695763	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695917	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695917 35696052	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695917 35696052 35696286	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer Cancer heart disease, cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695917 35696052 35696286 35696423	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695855 35695917 35696052 35696286 35696423 52644045	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695917 35696052 35696286 35696423	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695917 35696052 35696286 35696423 52644045 52644150	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus Cells)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line Chorionic villus	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer fertility, birth defects
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695855 35695855 35696052 35696052 35696286 35696423 52644045 52644150	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus Cells) NQH.11.3 (A549)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line Chorionic villus Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer fertility, birth defects Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695763 35695855 35695917 35696052 35696052 35696286 35696423 52644045 52644150 52644229 52644296	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus Cells) NQH.11.3 (A549) NQH.11.4 (U266B1)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line Chorionic villus Cancer Cell line Chore Cell line Chore Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer fertility, birth defects Cancer Cancer Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695855 35695855 35696052 35696052 35696286 35696423 52644045 52644150 52644229 52644296 52644332	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus Cells) NQH.11.3 (A549) NQH.11.4 (U266B1) NQH.11.5 (Daoy)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line Chorionic villus Cancer Cell line Chorionic villus Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer fertility, birth defects Cancer Cancer Cancer Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695855 35695855 35696052 35696052 35696423 52644045 52644150 52644229 52644296 52644332 52644507	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG-63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus Cells) NQH.11.3 (A549) NQH.11.4 (U266B1) NQH.11.5 (Daoy) NQH.11.6 (SW1783)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line Chorionic villus Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer Cancer fertility, birth defects Cancer Cancer Cancer Cancer Cancer Cancer Cancer Cancer Cancer Cancer Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695855 35695855 35696052 35696052 35696423 35696423 52644045 52644150 52644229 52644296 52644332 52644507 52645080	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG-63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus Cells) NQH.11.4 (U266B1) NQH.11.5 (Daoy) NQH.11.5 (Daoy) NQH.11.6 (SW1783) NQH.12.1 (U-118MG)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer Cancer Cancer feart disease, cancer heart disease, cancer fertility, birth defects Cancer Cancer Cancer Cancer Cancer Cancer Cancer Cancer Cancer Cancer Cancer Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695855 35695855 35696052 35696052 35696286 35696423 52644045 52644045 52644229 52644296 52644332 52644507 52645080 52645129	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG-63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus Cells) NQH.11.4 (U266B1) NQH.11.5 (Daoy) NQH.11.6 (SW1783) NQH.11.6 (SW1783) NQH.12.1 (U-118MG) NQH.12.2 (A204)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer fertility, birth defects Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695855 35695855 35696052 35696052 35696286 35696423 52644045 52644045 52644150 52644229 52644296 52644332 52644507 52645080 52645129 52645156	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus Cells) NQH.11.3 (A549) NQH.11.4 (U266B1) NQH.11.5 (Daoy) NQH.11.6 (SW1783) NQH.11.1 (U-118MG) NQH.12.2 (A204) NQH.12.3 (T24)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line Chorionic villus Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer fertility, birth defects Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695855 35695855 35696052 35696052 35696028 35696423 52644045 52644045 52644150 52644229 52644296 52644332 52644507 52645080 52645129 52645156 52646317	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG-63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus Cells) NQH.11.4 (U266B1) NQH.11.5 (Daoy) NQH.11.6 (SW1783) NQH.11.1 (U-118MG) NQH.12.2 (A204) NQH.12.3 (T24) NQH.12.4 (G-401)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line Chorionic villus Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer fertility, birth defects Cancer
29331827 29331828 29331830 87168518 87168559 35695763 35695855 35695855 35695855 35696052 35696052 35696423 52644045 52644045 52644229 52644296 52644332 52644507 52645080 52645129	NQH 8.5 (Spinal cord) NQH 8.6 (stomach) NQH 8.7 (Trachea) NQH 8.7 (Trachea) NQH 9.1 (Sized-MG- 63_treatment pool) NQH 9.2 (Sized-HEPG2 untreated) NQH.10.1 (MCF-7untreated) NQH.10.2 (U-937_treatment pool) NQH.10.3 (JAR) NQH.10.4 (PA-1) NQH.10.5 (CADMEC) NQH.10.6 (CADMEC_LA) NQH.11.1 (SK-PN-DW) NQH.11.2 (Chorionic Villus Cells) NQH.11.3 (A549) NQH.11.4 (U266B1) NQH.11.5 (Daoy) NQH.11.6 (SW1783) NQH.11.1 (U-118MG) NQH.12.2 (A204) NQH.12.3 (T24)	Spinal chord Stomach Trachea Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Endothelial cells Endothelial cells Cancer Cell line Chorionic villus Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line Cancer Cell line	paralysis, neurodegenerative disorders Stomach cancer Airway diseases, infection Cancer Cancer Cancer Cancer heart disease, cancer heart disease, cancer fertility, birth defects Cancer

60424179	NQH.14.1 (Yale75_breast carcinoma)	Breast carcinoma	Breast Cancer
60424269	NQH.14.2	Ovary tumor	Ovarian cancer
	(Yale78B_ovarytumor)		Ovarian Cancer
60431528	NQH.14.3	Prostate	Prostate Cancer
	(Yale79 prostateBPH)		1 Tostate Cancer
60431602	NQH.14.4	Prostate	Prostate Cancer
	(Yale80_ProstateAdenocarcin	1	1.00
	oma)		
60431735	NQH.14.5	Uterine Myorna	Uterine Cancer
	(Yale86_UterineMyoma)		
60431850	NQH.14.6	Myometrium	Fertility
	(Yale207_Myometrium)	<u> </u>	
60432049	NQH.15.1 (Yale99_cervix)	Cervix	Osteoporosis, cervical cancer
60432113	NQH.15.2		Hemophilia, Hypercoagulation, Idiopathic
İ	(Yale45_spleeniTP)		thrombocytopenic purpura, Immunodeficiencies, Graft
			vesus host
60432229	NQH.15.3 (Yale16_Skin)	Skin	wound healing, melanoma
60432289	NQH.15.4 (Yale137_Parotid)		
60433356	NQH.15.5	Small intestine	digestive diseases, obesity, diabetes
	(Yale38_SmallIntestine)		
60433438	NQH.15.6	Colon	Colon cancer
	(Yale28_ColonAscending)		
65274444	NQH.17.1 (Larynx)	Larynx	Cancer
65274572	NQH.17.2 (Duodenum)	Duodenum	
65274620	NQH.17.3 (Kidney, Primary		Diabetes, Autoimmune disease, Renal artery stenosis,
	tumors)		Interstitial nephritis, Glomerulonephritis, Polycystic
			kidney disease, Systemic lupus erythematosus, Renal
			tubular acidosis, IgA nephropathy, Hypercalceimia, Lesch- Nyhan syndrome
65274727	NQH.17.4 (Lung Pleura,	I are a	
032/4/2/	normal)	Lung	Airway diseases, infection
65274791	NQH.17.5 (Lung, Normal	Lung	Airway diseases, infection
İ	Adult)		· · · · · · · · · · · · · · · · · · ·
83373044	NQH.18.230 (Pooled adrenal	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy , Congenital Adrenal Hyperplasia,
	gland, placenta)	, 3	, , , , , , , , , , , , , , , , , , ,
85658542	NQH.18.560 (Pooled uterus,	Uterus	Infertility, birth defects
•	BeWo pool)		· ·
33656970	NQH.9.1 (MG-63_treatment	Cancer Cell line	Cancer
	pool)		
33657023	NQH.9.2 (HEPG2 untreated)		Von Hippel-Lindau (VHL) syndrome,
			Cirrhosis, Transplantation
33657084	NQH.9.3 (PC3_untreated)	Cancer Cell line	Cancer
33657109	NQH.9.4 (TF-1_TPA)	Cancer Cell line	Cancer
33657182	NQH.9.5 (TF-1_TPO)	Cancer Cell line	Cancer
3 365 7349	NQH.9.6 (TF-1_Hemin)	Cancer Cell line	Cancer
33657402	NQH.9.7 (HFDPC)	Cancer Cell line	Cancer
264259	NQH1 (Mixture of eight adult & two fetal tissues)		
264288	NQH2 (Ten tissues plus		
20-1266	lymphocyte control)		
264448	 	Bone Marrow	Hemonbilia humana andari
	(Done matow)	Done Martow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmume disease, allergies,
			immunodeficiencies, transplantation, Graft vesus host,
			. Tall vesus 1105L
265017	NQH4.1 (lymph node)	Lymph Node	Lymphedema , Allergies
	1		-7pitevenia , Ariergies

265018	NQH4.2 (fetal kidney)	Fetal Kidney	Diabetes, Autoimmune disease, Renal artery stenosis,
		,	Interstitial nephritis, Glomerulonephritis, Polycystic
			kidney disease, Systemic lupus erythematosus, Renal
			tubular acidosis, IgA nephropathy, Hypercalceimia, Lesch
		1	Nyhan syndrome
66712502	NQH4.2 (Sized)		.,,
265019	NQH4.3 (pituitary gland)		Von Hippel-Lindau (VHL) syndrome , Alzheimer's
	(,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		disease, Stroke, Tuberous scierosis, hypercalceimia,
			Parkinson's disease, Huntington's disease, Cerebral palsy,
			Epilepsy, Lesch-Nyhan syndrome, Multiple
			sclerosis, Ataxia-
		Į	telangiectasia, Leukodystrophies, Behavioral disorders,
			Addiction, Anxiety, Pain, Neuroprotection, Obesity
66714117	NQH4.3 (Sized)		The state of the s
265020	NQH4.4 (testis)	testis	Infertility, birth defects
265021	NQH4.5 (fetal liver)	Fetal Liver	Von Hippei-Lindau (VHL) syndrome,
			Cirrhosis, Transplantation
265022	NQH4.6 (thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
18108376	NQH5.1 (MCF-7)	Breast cancer	Breast Cancer
18108387	NQH5.2 (CCRF-CEM)	Cancer Cell line	Cancer
264952	NRL1: HPLC		
	FRACTIONATION OF RE-		
	LIG		
263971	Old BB3 Baits		
263969	Old BB5 Baits		
263975	ORFSEL		
263972	OTHER Baits		
263978	pGALORF		
264106	PPBAITS		
264088	QC-YA7		
264089	QC-YA8		
264102	Resequenced Interactors	•	
264369	RRH.1		
60170394	RRH.10.1 (MCF-7untreated)	Breast cancer	Breast Cancer
60170615	RRH.10.2 (U-937_treatment	Cancer Cell line	Cancer
	pool)	1	
60170831	RRH.10.3 (JAR)	Cancer Cell line	Cancer
60174639	RRH.11.8 (HeLa)	Cancer Cell line	Cancer
264113	пQEA Baits		
263973	RRQEA_B5 baits		
29146498	SRD 3.1 (SKMC)	Cancer Cell line	Cancer
29146499	SRD 3.2 (SKMC)	Cancer Cell line	Cancer
29147620	SRD 3.3 (RPTEC)	Cancer Cell line	Cancer
29148627	SRD 3.4 (HRCE)	Cancer Cell line	Cancer
29148629	SRD 3.6 (HRE)	Cancer Cell line	Cancer
29148784	SRD 3.7 (HRE)	Cancer Cell line	Cancer
55810764	SRD.7.1 (Lymph Node)	Lymph Node	Lymphedema, Allergies
5514150	SRD.7.2 (pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
55811150	(PEIOTOE)		, area area, parie, aution

55911576	CDD 2.4 (Dississes Class)	161:	
55811576	SRD.7.4 (Pituitary Gland)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalceimia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxiatelangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
5 581 1957	SRD.7.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
55812038	SRD.7.6 (Fetal Kidney)	Fetal kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalceimia, Lesch-Nyhan syndrome
56181562	SRD.8.1 (Lymph Node)	Lymph Node	Lymphedema, Allergies
56181686	SRD.8.2 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
56182181	SRD.8.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy , Congenital Adrenal Hyperplasia,
56182323	SRD.8.4 (Pituitary Gland)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalceimia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxiatelangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
56182435	SRD.8.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
56182575	SRD.8.6 (Fetal Kidney)		Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalceimia, Lesch- Nyhan syndrome
32833986	SRD4: HL adapter		
56526486	SRD5.1:rr fragments		
33109954	SRD5: long-RXRJ		
56994075	SRD9.1 (CS/SC)	Cancer Cell line	Cancer
263977	TSC Screen 1		

Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences

```
<210> 1
<211> 312
<212> DNA
<213> Homo sapiens
gtegacagtg teatgggege cegaceggeg ateggegttg egaqeacqtt eqqtaceqeq
agaggaateg gtacegegeg aegaggagga egggegggea geteggegae gaqtqqteeq
gegagenneg atcegtetge etceccacca cegatetegg cagecaggtt gtegaggatg
tegetggtgg agegetgete egegtetget tettgeteag eegtettgeg ggaetgggee
tgctcctcgg ccttggtcag ggcagcgtca atgtcggcga aatccgattc tgccttcggc
300
tcagagacgc gt
312
<210> 2
<211> 104
<212> PRT
<213> Homo sapiens
<400> 2
Val Asp Ser Val Met Gly Ala Arg Pro Ala Ile Gly Val Ala Ser Thr
                                    10
Phe Gly Thr Ala Arg Gly Ile Gly Thr Ala Arg Arg Gly Gly Arg Ala
                                25
Gly Ser Ser Ala Thr Ser Gly Pro Ala Ser Xaa Asp Pro Ser Ala Ser
                            40
Pro Pro Pro Ile Ser Ala Ala Arg Leu Ser Arg Met Ser Leu Val Glu
                        55
Arg Cys Ser Ala Ser Ala Ser Cys Ser Ala Val Leu Arg Asp Trp Ala
                                        75
Cys Ser Ser Ala Leu Val Arg Ala Ala Ser Met Ser Ala Lys Ser Asp
                                    90
Ser Ala Phe Gly Ser Glu Thr Arg
            100
```

```
<210> 3
<211> 987
<212> DNA
<213> Homo sapiens
<400> 3
gageteggta ecceegtegt egeegteaag gegeaggtea aagtgggtgg eegtggeaag
gctggcggtg taaagatcgc taagtccccg gccgaggcgg ctgaacgcgc cgaggccatc
ctcggtatgg acatcaaggg tcacaccgtt cacaaggtga tggtcgctga aggtgctgac
attgccgagg aatactactt ctcgatcttg ctcgaccgtg gagagcgtcg ctaccttgcg
atgtgctcgc gtgagggtgg catggacatc gagacccttg ctaaggagcg ccccgaggct
ctcgccaagg tgccggtcga cccgattgac ggcgttgacg atgctaaagc ccgcgaaatc
cttagegagg cagggtteec egactetgag caggaegeta tegteeegge tgtteteaag
420
ctgtgggaga cctaccgtga cgaggatgcc accetcgtcg aggtcaaccc gatgatcaag
480
accggcgatg gacgcatcct ggctatcgac ggcaagatga ctgttgacaa caacgcatcc
ttccgccagc ctgaccgcgc cggcttggtg gatcgcgcca ccaccgaccc gctcgagttg
cgtgccggcg agctcggtct caactacgtc aagcttgacg gcaacgtcgg cgtcatcgga
660
aacggtgctg gtctggtcat gagcaccctt gactgcgttg cgtacgccgg tgagaacttc
720
cogggatoto cagotocogo taacttooto gacatoggtg goggegeoto ggoogagato
atggccaacg gccttgacct catcatgagt gacgagcagg tgcgttccgt gttcgtcaac
gtctttggcg gtatcaccgc ctgtgaccag gtggcgcttg gtatcaaggg cgctctcgag
aagttgggcg acaaggccgt gaagccgctc gtcgtccgtc tggacggcaa cgctgtggcc
gagggcagaa agattctcga ggaattc
987
<210> 4
<211> 329
<212> PRT
<213> Homo sapiens
<400> 4
Glu Leu Gly Thr Pro Val Val Ala Val Lys Ala Gln Val Lys Val Gly
                                    10
1
Gly Arg Gly Lys Ala Gly Gly Val Lys Ile Ala Lys Ser Pro Ala Glu
            20
Ala Ala Glu Arg Ala Glu Ala Ile Leu Gly Met Asp Ile Lys Gly His
                            40
        35
```

```
Thr Val His Lys Val Met Val Ala Glu Gly Ala Asp Ile Ala Glu Glu
                       55
Tyr Tyr Phe Ser Ile Leu Leu Asp Arg Gly Glu Arg Arg Tyr Leu Ala
                                       75
                   70
Met Cys Ser Arg Glu Gly Gly Met Asp Ile Glu Thr Leu Ala Lys Glu
                                  90
Arg Pro Glu Ala Leu Ala Lys Val Pro Val Asp Pro Ile Asp Gly Val
                              105
Asp Asp Ala Lys Ala Arg Glu Ile Leu Ser Glu Ala Gly Phe Pro Asp
                                             125
                          120
Ser Glu Gln Asp Ala Ile Val Pro Ala Val Leu Lys Leu Trp Glu Thr
                      135
                                          140
Tyr Arg Asp Glu Asp Ala Thr Leu Val Glu Val Asn Pro Met Ile Lys
                                      155 160
                  150
Thr Gly Asp Gly Arg Ile Leu Ala Ile Asp Gly Lys Met Thr Val Asp
                                  170
               165
Asn Asn Ala Ser Phe Arg Gln Pro Asp Arg Ala Gly Leu Val Asp Arg
                              185
Ala Thr Thr Asp Pro Leu Glu Leu Arg Ala Gly Glu Leu Gly Leu Asn
       195
                          200
Tyr Val Lys Leu Asp Gly Asn Val Gly Val Ile Gly Asn Gly Ala Gly
                      215
                                          220
Leu Val Met Ser Thr Leu Asp Cys Val Ala Tyr Ala Gly Glu Asn Phe
                  230
                                      235
Pro Gly Ser Pro Ala Pro Ala Asn Phe Leu Asp Ile Gly Gly Ala
               245
                                  250
Ser Ala Glu Ile Met Ala Asn Gly Leu Asp Leu Ile Met Ser Asp Glu
                              265
Gln Val Arg Ser Val Phe Val Asn Val Phe Gly Gly Ile Thr Ala Cys
                           280
       275
Asp Gln Val Ala Leu Gly Ile Lys Gly Ala Leu Glu Lys Leu Gly Asp
                       295
                                          300
Lys Ala Val Lys Pro Leu Val Val Arg Leu Asp Gly Asn Ala Val Ala
                   310
                                       315
Glu Gly Arg Lys Ile Leu Glu Glu Phe
               325
<210> 5
<211> 622
<212> DNA
<213> Homo sapiens
<400> 5
acgcgtggcc tagacctgaa tccattcctc acaaaacagt ctccctcctt gccatgggat
tccctggggt tttcacaggc cacagctcta atggtctgca gcaggttacc ttgttcccca
qaacataget tgtcataaca tetetgeagg gtteteecaa acceetttet geetggeaac
agetgacate acacetaget gtaagteeet gtagategea aattactttt tggagaetgg
gggtagcagg ggcattgggg taatagcctt ctagcccttt ttgagggaaa cacatgggtg
```

```
aggetatttt ggggetggga agtgggggee tggtgteece tggatggetg tgetggeete
360
tggctgcaag ggagagggc acaggcaagg acatgacccc cgtcaaccct gagccccctc
420
cagaaattta accagageet gteeeteett tettgeetge ecceaacate teacaateee
480
teetgtgatg geagatgtet ecatetaete tacagacace tgeaactate atteeettga
teegtggtaa ttaggaggga acteetetgt gaagaacege ttetaceate etetttaga
aactetttet ceaetgggat ce
622
<210> 6
<211> 121
<212> PRT
<213> Homo sapiens
<400> 6
Met Ser Leu Pro Val Pro Leu Ser Leu Ala Ala Arg Gly Gln His Ser
                 5
                                    10
His Pro Gly Asp Thr Arg Pro Pro Leu Pro Ser Pro Lys Ile Ala Ser
                                25
Pro Met Cys Phe Pro Gln Lys Gly Leu Glu Gly Tyr Tyr Pro Asn Ala
                            40
                                                45
       35
Pro Ala Thr Pro Ser Leu Gln Lys Val Ile Cys Asp Leu Gln Gly Leu
                                            60
                        55
Thr Ala Arg Cys Asp Val Ser Cys Cys Gln Ala Glu Arg Gly Leu Gly
                                        75
                    70
Glu Pro Cys Arg Asp Val Met Thr Ser Tyr Val Leu Gly Asn Lys Val
                                    90
Thr Cys Cys Arg Pro Leu Glu Leu Trp Pro Val Lys Thr Pro Gly Asn
                                105
           100
Pro Met Ala Arg Arg Glu Thr Val Leu
<210> 7
<211> 480
<212> DNA
<213> Homo sapiens
<400> 7
agatetttte agaageggtg ggtgggeegt teatateeae gtgaatggge eeaaaaggea
agaaaaccca cacctctcct ccctcccca gtggctgtgg ctttcctagg gacaatagga
tgaatgggct ttcagtgtgg ggacagcaaa acatgcacta gggcccagag tggcagttct
cttqqtqtqg agagtgcctg ccacaggcct tggccagagc ccgtgaggga gtggtgtgtg
aaaggccacc tccacgtggg taagcgtgag gacttggact tctctggcac tgagatggga
cetectgeet gtgggagtea tetggeeace accetgggge cagtaaaggt tggagetaga
360
```

```
agggtcgtcc tccctgactt gagctctgag ggctttgcct gcccagccag agcggcaagg
420
cacaqqqqac cctcqqqqac gcccatqqcc accctqqqqa agacaqqqct cctcacqcqt
480
<210> 8
<211> 119
<212> PRT
<213> Homo sapiens
<400> 8
Met Gly Phe Gln Cys Gly Asp Ser Lys Thr Cys Thr Arg Ala Gln Ser
                                    10
Gly Ser Ser Leu Gly Val Glu Ser Ala Cys His Arg Pro Trp Pro Glu
            20
                                25
Pro Val Arg Glu Trp Cys Val Lys Gly His Leu His Val Gly Lys Arg
                            40
Glu Asp Leu Asp Phe Ser Gly Thr Glu Met Gly Pro Pro Ala Cys Gly
                        55
Ser His Leu Ala Thr Thr Leu Gly Pro Val Lys Val Gly Ala Arg Arg
                    70
Val Val Leu Pro Asp Leu Ser Ser Glu Gly Phe Ala Cys Pro Ala Arg
                                    90
                85
Ala Ala Arg His Arg Gly Pro Ser Gly Thr Pro Met Ala Thr Leu Gly
           100
                                105
Lys Thr Gly Leu Leu Thr Arg
        115
<210> 9
<211> 428
<212> DNA
<213> Homo sapiens
tccqqaaatt ttccatgaga gactggggca tcgagcagaa gtggatgtct gttctcctgc
60
ctctgctgct actttacaat gatccgttct tccccctctc cttcctggtc aacagctggc
120
teccagggat getggatgae etettteagt ceatgtteet gtgegeeetg etgetettet
ggctgtgcgt gtaccacggg attcgtgtcc agggagaaag aaagtgttta actttctatt
tgcctaaatt cttcattgtt ggactattgt ggttggcttc tgttacgcta ggaatatggc
aaacagttaa cgaattacat gatccaatgt accagtatcg agttgatacc ggaaattttc
agggaatgaa ggtcttcttc atggtggtgg cagcggtgta cattctgtac ctcttgttct
420
tgatagtg
428
<210> 10
<211> 138
<212> PRT
```

<213> Homo sapiens <400> 10 Met Arg Asp Trp Gly Ile Glu Gln Lys Trp Met Ser Val Leu Leu Pro 10 Leu Leu Leu Tyr Asn Asp Pro Phe Phe Pro Leu Ser Phe Leu Val Asn Ser Trp Leu Pro Gly Met Leu Asp Asp Leu Phe Gln Ser Met Phe 40 Leu Cys Ala Leu Leu Leu Phe Trp Leu Cys Val Tyr His Gly Ile Arg 55 Val Gln Gly Glu Arg Lys Cys Leu Thr Phe Tyr Leu Pro Lys Phe Phe 75 Ile Val Gly Leu Leu Trp Leu Ala Ser Val Thr Leu Gly Ile Trp Gln 85 90 Thr Val Asn Glu Leu His Asp Pro Met Tyr Gln Tyr Arg Val Asp Thr 100 105 Gly Asn Phe Gln Gly Met Lys Val Phe Phe Met Val Val Ala Ala Val 120 125 Tyr Ile Leu Tyr Leu Leu Phe Leu Ile Val <210> 11 <211> 453 <212> DNA <213> Homo sapiens cttaaqaatc qcctcactca acqqtcaqct tqccqaccat qcccgcctga taatgccccg gaatgttgca ggcaaactca agaccggtgg ccttggtgaa ggtccaggtc agctcggcgg acttgcccgg ctcgaccagc acgctgttgg ggtcgtcatg cttcatgccg cccatatcgc catgecccat ggeggegtgg tecatettge ceatgeeggt ggeegtgage atgeegetgg cttgcatctt gagcatttct ttctggtgtt cggcgtgcat cgccgcatca cccagattga attegtgeag taactggeet ttgttgacca geacaaageg caeggtetea eeggetttta catccagage cttgggcgaa aaggaaatgt cctgcagggt gacttccacg gtgcgcgtgg ctttatcggc cggtgccggg tggccaaacg cgt 453 <210> 12 <211> 130 <212> PRT <213> Homo sapiens Met Leu Gln Ala Asn Ser Arg Pro Val Ala Leu Val Lys Val Gln Val Ser Ser Ala Asp Leu Pro Gly Ser Thr Ser Thr Leu Leu Gly Ser Ser

```
30
           20
Cys Phe Met Pro Pro Ile Ser Pro Cys Pro Met Ala Ala Trp Ser Ile
                           40
Leu Pro Met Pro Val Ala Val Ser Met Pro Leu Ala Cys Ile Leu Ser
Ile Ser Phe Trp Cys Ser Ala Cys Ile Ala Ala Ser Pro Arg Leu Asn
Ser Cys Ser Asn Trp Pro Leu Leu Thr Ser Thr Lys Arg Thr Val Ser
               85
Pro Ala Phe Thr Ser Arg Ala Leu Gly Glu Lys Glu Met Ser Cys Arg
                               105
Val Thr Ser Thr Val Arg Val Ala Leu Ser Ala Gly Ala Gly Trp Pro
                           120
Asn Ala
   130
<210> 13
<211> 2034
<212> DNA
<213> Homo sapiens
<400> 13
nacgcgttcg gcgtagtccc cttcctgcca tccgtcctga gctccctgct gcccgtgctg
ggcgtggcca agcaggacac ggtgcgcgtg gccttctgct ccggggacct gcggctcctc
tgggccctac catgctggca ttttcctcca tgtgtcaaac acatgggttc agccagcgaa
gattecatgg gaceteeteg tgtgggaegt gtgeteecea ccacaaatgg aacgtteeet
gtttgcatct ggaggggttg gtggtcctgc tggctggagc agcctggggc cagaggaagc
cgtatcaacc ggctctgcag cgcttcagcg agggtgccct ggagtaccta gccaacctgg
accqagecce agaceceacg greaggaagg acgeetttge cacegacate treagegeet
420
acgatgttct cttccatcag tggctgcaga gtcgagaagc caagctccgt cttgccgtgg
480
tggaggetet ggggeetatg agecatetge tgeecagtga gaggetggaa gageagetge
ccaageteet ecctgggatt etegecetet acaagaagea egcagagace ttetaettgt
ccaagageet gggecagate etegaggeag etgtgagtgt gggeageege acaetggaga
cccagctgga tgccctcttg gctgcactgc actcccagat ctgtgtgcct gtggagtcct
caageceet ggtgatgagt aaccagaagg aggtgetgeg etgetteact gtgetggeet
quaqueteque tgacequeta etggeettee tgetgeccag getggacace accaatgaga
tggaagataa aaagcccttt atcctgtctt ccatgaggct tcctctcctg aacaccaaca
```

gcaaggtgaa gcgggcagtg gtgcaggtga ttagcgccca tggcccacca cggctacctg

1020

```
gagcageetg gaggtgagge gatgategag taaategtge ageagtgege getgeeeeee
1080
gagcaggagc ctgagaagcc aggccccggc agcaaggacc ccaaggccga cagcgtgcgg
gecateageg tgegeaceet etacetggte ageaceaceg tggacaggat gagteacgte
ctetggccat acctgctcca gttcctcacc cetgtgcgct tcactggggc cctgactccg
ctctgcagga gcctcgtgca tctggcgcag aagaggcagg aggccggggc cgacgccttc
ctcatccagt acgacgccca tgcgagcctc ccgtctccct atgctgtaac cggaagactg
ttggttgtgt cttccagccc ctacctaggg gacggacgtg gggcagcggc gctgcgcctc
ctcagtgttc tgcacccaaa cattcaccct ttgctgggtc agcattggga aacgactgtc
ccgctgctgc tggggtacct ggatgagcac acagaagaga ccctgccaca ggaggagtgg
gaggagaagc tgttgatggt gagggccggg gtacggccca tcctgggcct taaggtgttg
tetggeetgg ggggtgetgg ggtggeagag getgggeeae etgeetegae eteaeetegt
ggtttggetg gggagecaag gatcaggeag cateaagget gaagaeeeca geageettge
agegggggee ttgetgtgae aaggeaeegg eeetetagea gtegeageee eaagegtegg
1800
gggcaacete teaceetgee tggtgageea actgtggcat ggetgteece tgagggttgg
1860
ctctgccgcc cccggcctcc gctggaaggc ggtctgcagc ccctgcagcc acagcacatg
1920
gggatgtgcc caggetecag ecagecetgt gaggggtegg geteceagee eeteagtgge
atcttggcct gcagttcctg cgagacaccc tggccatcat ttctgacaac gcgt
2034
<210> 14
<211> 222
<212> PRT
<213> Homo sapiens
<400> 14
Ile Val Gln Gln Cys Ala Leu Pro Pro Glu Gln Glu Pro Glu Lys Pro
                                    10
Gly Pro Gly Ser Lys Asp Pro Lys Ala Asp Ser Val Arg Ala Ile Ser
Val Arg Thr Leu Tyr Leu Val Ser Thr Thr Val Asp Arg Met Ser His
                            40
Val Leu Trp Pro Tyr Leu Leu Gln Phe Leu Thr Pro Val Arg Phe Thr
                                             60
    50
Gly Ala Leu Thr Pro Leu Cys Arg Ser Leu Val His Leu Ala Gln Lys
                                        75
Arg Gln Glu Ala Gly Ala Asp Ala Phe Leu Ile Gln Tyr Asp Ala His
```

```
90
               85
Ala Ser Leu Pro Ser Pro Tyr Ala Val Thr Gly Arg Leu Leu Val Val
                                                    110
           100
                                105
Ser Ser Ser Pro Tyr Leu Gly Asp Gly Arg Gly Ala Ala Leu Arg
                            120
                                                125
Leu Leu Ser Val Leu His Pro Asn Ile His Pro Leu Leu Gly Gln His
                                            140
                        135
Trp Glu Thr Thr Val Pro Leu Leu Leu Gly Tyr Leu Asp Glu His Thr
                                        155
                    150
Glu Glu Thr Leu Pro Gln Glu Glu Trp Glu Glu Lys Leu Leu Met Val
                                    170
Arg Ala Gly Val Arg Pro Ile Leu Gly Leu Lys Val Leu Ser Gly Leu
                                                    190
                                185
            180
Gly Gly Ala Gly Val Ala Glu Ala Gly Pro Pro Ala Ser Thr Ser Pro
                            200
       195
Arg Gly Leu Ala Gly Glu Pro Arg Ile Arg Gln His Gln Gly
                        215
<210> 15
<211> 363
<212> DNA
<213> Homo sapiens
<400> 15
nacgcgttgc tggctcgcca cggcaagggc catgtcggct gcgatatctg caagccggcg
gtgggttcga tccttgcctc gtgctggaac cagccgatca tggacccggc gttggtgccg
ttgcaggaca ccaatgacac cttcatggcc aacatgcaga agaacggtac ctattcgatc
atcccgcgta tcgccggcgg cgagatcacc ccggacaaac tgatcgccct cggcgcggtg
gegaagaaat acgatetgta caccaagate accggeggee ageggatega cetgttegge
geccagttge acgaattgee geagatetgg ggegagetgg tggatgeegg attegagaee
360
ggt
363
<210> 16
<211> 121
<212> PRT
<213> Homo sapiens
<400> 16
Xaa Ala Leu Leu Ala Arg His Gly Lys Gly His Val Gly Cys Asp Ile
                                    10
Cys Lys Pro Ala Val Gly Ser Ile Leu Ala Ser Cys Trp Asn Gln Pro
                                25
Ile Met Asp Pro Ala Leu Val Pro Leu Gln Asp Thr Asn Asp Thr Phe
                            40
Met Ala Asn Met Gln Lys Asn Gly Thr Tyr Ser Ile Ile Pro Arg Ile
Ala Gly Gly Glu Ile Thr Pro Asp Lys Leu Ile Ala Leu Gly Ala Val
```

```
70
                                        75
65
Ala Lys Lys Tyr Asp Leu Tyr Thr Lys Ile Thr Gly Gly Gln Arg Ile
                                    90
                85
Asp Leu Phe Gly Ala Gln Leu His Glu Leu Pro Gln Ile Trp Gly Glu
                                105
            100
Leu Val Asp Ala Gly Phe Glu Thr Gly
<210> 17
<211> 682
<212> DNA
<213> Homo sapiens
<400> 17
gaattccatt ttgtggagta agaggtgact ggggtatagg gtacaaccca tagccatcca
tgttcatctt tgttttgaat ataattggct agaagatata catatatcta tgtaacttcc
120
totagcatoc tocagtatgg aggotgcatt aagactgcat gaaggagagg gagagaaggg
aqaaacaqaq caqctqqaca aqaqqacaqg tatagggaat aagggagaag ccagtaaggc
aggaaagacc ctccgtgaca aaggggcagg gaacagaact caaacattta atggcaggta
acccaggtta gaatggtaaa ttgaaaggtg aatataaagg gagaatggtg aaatgaattt
totgaaatta attgotgtgt ttatagtttt tagocatgca toggaatcac otcaggacto
cacteceaat caattatata tetgggggag gaccaaggeg ttggtatttt teagaagete
cactggtgat tetgacagea cagetaggat taagaaactg atcaatggga acageatgee
tgttgcagag gagcttccct gggaaatgtc acacacagaa catcaatctt ccttccccac
600
tectgagate ceteattett tggcaccagg aacagttgca attagtaaac cetggtteec
toctotctca caaatcocaa ga
682
<210> 18
<211> 110
<212> PRT
<213> Homo sapiens
<400> 18
Met Asn Phe Leu Lys Leu Ile Ala Val Phe Ile Val Phe Ser His Ala
Ser Glu Ser Pro Gln Asp Ser Thr Pro Asn Gln Leu Tyr Ile Trp Gly
Arg Thr Lys Ala Leu Val Phe Phe Arg Ser Ser Thr Gly Asp Ser Asp
                            40
Ser Thr Ala Arg Ile Lys Lys Leu Ile Asn Gly Asn Ser Met Pro Val
    50
                        55
                                            60
Ala Glu Glu Leu Pro Trp Glu Met Ser His Thr Glu His Gln Ser Ser
```

```
65
                    70
                                        75
Phe Pro Thr Pro Glu Ile Pro His Ser Leu Ala Pro Gly Thr Val Ala
                                    90
Ile Ser Lys Pro Trp Phe Pro Ala Val Ser Gln Ile Ala Arg
                                105
<210> 19
<211> 515
<212> DNA
<213> Homo sapiens
<400> 19
cttggctggc agacatggga cctgcttccc tcttacaccc cagtcttggc aaggatcatg
cocccatcte aactatgtta gecagtetgg etgtteaett agteactaca gtttgettet
cgtctgcagt gcagtcttgg gctataagaa acactgggcc actcaatacc tcccccttt
tggcccttct cctcctctgg tccatgggtg gggttggggg gagcccagtt tcagcaccag
cagctggagc ccataccaca ctcatttttc agttctggct gtgggagccc ctcccacagg
tttcagttcc ccaagcccca ggcctgagtt ttttttattg caaaagctgg ttgttgttgt
ggctagctcc caggcgtgtg aggtgcagct tqctaagtaa gaqctaggaa aqaqaatagg
qtcctqctgt aggtgtccag tctgaaqgaa tqcctqgqat acttcctcaa qcaqttcctt
ctcacagtct cctggctgct ccgcatgtca gatct
515
<210> 20
<211> 130
<212> PRT
<213> Homo sapiens
<400> 20
Met Gly Pro Ala Ser Leu Leu His Pro Ser Leu Gly Lys Asp His Ala
                                    10
Pro Ile Ser Thr Met Leu Ala Ser Leu Ala Val His Leu Val Thr Thr
                                25
Val Cys Phe Ser Ser Ala Val Gln Ser Trp Ala Ile Arg Asn Thr Gly
                            40
Pro Leu Asn Thr Ser Pro Leu Leu Ala Leu Leu Leu Trp Ser Met
Gly Gly Val Gly Gly Ser Pro Val Ser Ala Pro Ala Ala Gly Ala His
Thr Thr Leu Ile Phe Gln Phe Trp Leu Trp Glu Pro Leu Pro Gln Val
                85
                                    90
Ser Val Pro Gln Ala Pro Gly Leu Ser Phe Phe Tyr Cys Lys Ser Trp
            100
                               105
Leu Leu Trp Leu Ala Pro Arg Arg Val Arg Cys Ser Leu Leu Ser
                            120
Lys Ser
```

130 <210> 21 <211> 390 <212> DNA <213> Homo sapiens <400> 21 gtgcgcacaa aagagcacgt tcgcaagggg aggaagagcg tgccaccggt tctgccgagc tagacgcggt gcctatggtt gcggaggacc atggagtgaa gcgagtaaga ctagatgatg 120 caacaaatgt gcctgagggt gaaatggcac gagccagtgc caatgagggc atgacacctg 180 ttaaccacga caaataccct tetgteettt taaatgaage ggeecagget teattactgg atacaatgac tgcttgcact gatgggttca caattgagca attggagctt acacgatctc tatgttatga aagagtatta gcacatcgat cctcatggga tcgttcagcc ctggctcaag aattaaagca agttgtccaa ggcatccatn 390 <210> 22 <211> 105 <212> PRT <213> Homo sapiens <400> 22 Met Val Ala Glu Asp His Gly Val Lys Arg Val Arg Leu Asp Asp Ala 1.0 Thr Asn Val Pro Glu Gly Glu Met Ala Arg Ala Ser Ala Asn Glu Gly 25 Met Thr Pro Val Asn His Asp Lys Tyr Pro Ser Val Leu Leu Asn Glu 40 Ala Ala Gln Ala Ser Leu Leu Asp Thr Met Thr Ala Cys Thr Asp Gly 55 60 Phe Thr Ile Glu Gln Leu Glu Leu Thr Arg Ser Leu Cys Tyr Glu Arg 70 75 Val Leu Ala His Arg Ser Ser Trp Asp Arg Ser Ala Leu Ala Gln Glu Leu Lys Gln Val Val Gln Gly Ile His 100 <210> 23 <211> 385 <212> DNA <213> Homo sapiens <400> 23 nteteggagg cegacageet ggegggetgg aageeetegg tgtaceaegt getgeteate etgggeetgt tegeegtget getgteetge tgegeetegg ecatgtacae cagegtggag

```
ggetgggaet aegtggaete getetaette tgettegtea eetteageae eateggette
ggggacctgg tgagcagcca gcacgccgcc taccggaacc aggggctcta ccgcctgggc
aactteetet teateetget eggegtgtge tgeatttaet egetetteaa egteatetee
atcctcatca agcaggtgct caactggatg ctgcgcaagc tgagctgccg ctgctgcgcg
cgctgctgcc cggctcctgg cgcgc
385
<210> 24
<211> 128
<212> PRT
<213> Homo sapiens
<400> 24
Xaa Ser Glu Ala Asp Ser Leu Ala Gly Trp Lys Pro Ser Val Tyr His
                                    10
                                                         15
Val Leu Leu Ile Leu Gly Leu Phe Ala Val Leu Leu Ser Cys Cys Ala
                                25
Ser Ala Met Tyr Thr Ser Val Glu Gly Trp Asp Tyr Val Asp Ser Leu
                                                45
                            40
Tyr Phe Cys Phe Val Thr Phe Ser Thr Ile Gly Phe Gly Asp Leu Val
                        55
Ser Ser Gln His Ala Ala Tyr Arg Asn Gln Gly Leu Tyr Arg Leu Gly
                                        75
Asn Phe Leu Phe Ile Leu Leu Gly Val Cys Cys Ile Tyr Ser Leu Phe
                85
Asn Val Ile Ser Ile Leu Ile Lys Gln Val Leu Asn Trp Met Leu Arg
                                105
                                                    110
           100
Lys Leu Ser Cys Arg Cys Cys Ala Arg Cys Cys Pro Ala Pro Gly Ala
                            120
                                                125
<210> 25
<211> 337
<212> DNA
<213> Homo sapiens
<400> 25
ccatgggaga gaccgtgcat titettetag gtetgegtgg gaagteactg cagagttteg
aggagggag ttcccagctc tgtatttttg aagggtcagt cttgttgctt ggaccagtga
120
ggagccccgt gggatccaga ctcgagtggg tggagccggg gcaggtggga gcagagacac
tggaggaaag ctggtcgaat gcactgtgta tttggaggca gaaccagcag agggtcctct
gggttgagtg tagggcaaaa gagaaagaag gcaccaagcc tggggtctgg gttttctctc
ttacacttgc tgggtggacg gtggtgccac tgaatga
337
<210> 26
```

```
<211> 111
<212> PRT
<213> Homo sapiens
<400> 26
Met Gly Glu Thr Val His Phe Leu Leu Gly Leu Arg Gly Lys Ser Leu
Gln Ser Phe Glu Glu Gly Ser Ser Gln Leu Cys Ile Phe Glu Gly Ser
                                25
Val Leu Leu Gly Pro Val Arg Ser Pro Val Gly Ser Arg Leu Glu
                            40
Trp Val Glu Pro Gly Gln Val Gly Ala Glu Thr Leu Glu Glu Ser Trp
Ser Asn Ala Leu Cys Ile Trp Arg Gln Asn Gln Gln Arg Val Leu Trp
                    70
                                        75
Val Glu Cys Arg Ala Lys Glu Lys Glu Gly Thr Lys Pro Gly Val Trp
Val Phe Ser Leu Thr Leu Ala Gly Trp Thr Val Val Pro Leu Asn
                               105
<210> 27
<211> 333
<212> DNA
<213> Homo sapiens
<400> 27
ccgacgtcga atatccatgc agccgcgccg aggatggaga gagcgatgga gcaactcaac
cgcctgacgc gctcgctgcg ccgcgcgcgc accgtggagt tgcccgagga taatgaaact
getgtttata cattaatgcc aatggttatg getgatcaac acaggtetgt ttetgaacta
ctatcaaatt caaaatttga tgtcaattat gcattcggac gtgtgaaaag aagcttgctt
cacattgcag caaattgtgg atcggtggaa tgcttggttt tgctgttaaa gaaaggagca
aatcctaact atcaagatat ttcaggctgt aca
333
<210> 28
<211> 111
<212> PRT
<213> Homo sapiens
Pro Thr Ser Asn Ile His Ala Ala Pro Arg Met Glu Arg Ala Met
                                    10
Glu Gln Leu Asn Arg Leu Thr Arg Ser Leu Arg Arg Ala Arg Thr Val
                                25
Glu Leu Pro Glu Asp Asn Glu Thr Ala Val Tyr Thr Leu Met Pro Met
                            40
Val Met Ala Asp Gln His Arg Ser Val Ser Glu Leu Leu Ser Asn Ser
Lys Phe Asp Val Asn Tyr Ala Phe Gly Arg Val Lys Arg Ser Leu Leu
```

```
70
His Ile Ala Ala Asn Cys Gly Ser Val Glu Cys Leu Val Leu Leu Leu
                                    90
               85
Lys Lys Gly Ala Asn Pro Asn Tyr Gln Asp Ile Ser Gly Cys Thr
                                105
<210> 29
<211> 375
<212> DNA
<213> Homo sapiens
<400> 29
negcegteeg tgetggetat tatgaeggeg ggtagegaec agggegagga ggteaacteg
gagagetatt tgagegeegt gaegeegetg agteecaaag agattegtea getgeeeege
tacaatatca cgatcaagcg cgtcgtgaac atgacgggca agggccgcac gccgagctgg
tactogotog togtggotgg caatggtogg ggootogtgg gotatggoga aggcaaagat
actaacatca geegegegaa caaaaaggeg ttecaegeeg eggtgaaaaa catggaettg
gtateggtee accggtegaa gagtggegee aacacgeteg ageeeceegt egagggeege
tggggcgcta cgcgt
375
<210> 30
<211> 125
<212> PRT
<213> Homo sapiens
<400> 30
Xaa Pro Ser Val Leu Ala Ile Met Thr Ala Gly Ser Asp Gln Gly Glu
Glu Val Asn Ser Glu Ser Tyr Leu Ser Ala Val Thr Pro Leu Ser Pro
           20
                                25
Lys Glu Ile Arg Gln Leu Pro Arg Tyr Asn Ile Thr Ile Lys Arg Val
                            40
Val Asn Met Thr Gly Lys Gly Arg Thr Pro Ser Trp Tyr Ser Leu Val
                                            60
Val Ala Gly Asn Gly Arg Gly Leu Val Gly Tyr Gly Glu Gly Lys Asp
Thr Asn Ile Ser Arg Ala Asn Lys Lys Ala Phe His Ala Ala Val Lys
                                    90
               85
Asn Met Asp Leu Val Ser Val His Arg Ser Lys Ser Gly Ala Asn Thr
                                105
Leu Glu Pro Pro Val Glu Gly Arg Trp Gly Ala Thr Arg
                            120
<210> 31
<211> 375
<212> DNA
<213> Homo sapiens
```

```
<400> 31
accegetette geeteagett teetetegaaa tegaagteeg teecaaaagt gegegaagage
gggagcaggc acttacgagc ctgcgcgtca gggatgcttc ctgggcccct gagagtgcag
120
agatteetgg atecagaget geggetggge ggetgeaget gegeetggga gtgeaggget
180
cccgccctgc cagctcaaaa ggaaatgggg gctcctgcct gttcctggct cctgttggcc
ctgcagagtg cacaaaccta gccgcgcttc ctccactgca gcttacgtct ttgcagcagc
cactecegat gggetgecae tgecatetgt gagaccataa tgtgtgeaat ttgagactea
tggcctgcat tgttt
375
<210> 32
<211> 118
<212> PRT
<213> Homo sapiens
<400> 32
Met Gln Ala Met Ser Leu Lys Leu His Thr Leu Trp Ser His Arg Trp
                 5
                                    10
1
Gln Trp Gln Pro Ile Gly Ser Gly Cys Cys Lys Asp Val Ser Cys Ser
            20
Gly Gly Ser Ala Ala Arg Phe Val His Ser Ala Gly Pro Thr Gly Ala
        35
                            40
                                                 45
Arg Asn Arg Gln Glu Pro Pro Phe Pro Phe Glu Leu Ala Gly Arg Glu
                        55
Pro Cys Thr Pro Arg Arg Ser Cys Ser Arg Pro Ala Ala Ala Leu Asp
                    70
                                        75
Pro Gly Ile Ser Ala Leu Ser Gly Ala Gln Glu Ala Ser Leu Thr Arg
                                    90
Arg Leu Val Ser Ala Cys Ser Arg Ser Ser Pro Leu Leu Ala Pro Thr
            100
                                105
Ser Ile Ser Glu Gln Ser
       115
<210> 33
<211> 351
<212> DNA
<213> Homo sapiens
<400> 33
ccatqcaqcc caaccgttgg cgataaagtc cgtttaggcg ataccaattt atgggcaacc
attgaacaag atttattaac caaaggtgat gagtgtaaat ttggtggcgg taaaagtgtg
cgtgatggta tggcgcaaag cggcaccgca actcgcgaca atccaaatgt attggatttt
qtgattacca atgtgatgat cattgatgcc aaattaggca ttatcaaagc cgatattggt
240
```

```
atteqeqatq qteqtattqt eqqtateqqa caaqeaqqta accetgacae catggatgae
gtcacgccaa acatgattat cggtgctagc acagaagtac ataacggtgc a
351
<210> 34
<211> 117
<212> PRT
<213> Homo sapiens
<400> 34
Pro Cys Ser Pro Thr Val Gly Asp Lys Val Arg Leu Gly Asp Thr Asn
                                    10
Leu Trp Ala Thr Ile Glu Gln Asp Leu Leu Thr Lys Gly Asp Glu Cys
Lys Phe Gly Gly Lys Ser Val Arg Asp Gly Met Ala Gln Ser Gly
Thr Ala Thr Arg Asp Asn Pro Asn Val Leu Asp Phe Val Ile Thr Asn
Val Met Ile Ile Asp Ala Lys Leu Gly Ile Ile Lys Ala Asp Ile Gly
Ile Arg Asp Gly Arg Ile Val Gly Ile Gly Gln Ala Gly Asn Pro Asp
Thr Met Asp Asp Val Thr Pro Asn Met Ile Ile Gly Ala Ser Thr Glu
                                105
                                                    110
            100
Val His Asn Gly Ala
       115
<210> 35
<211> 355
<212> DNA
<213> Homo sapiens
<400> 35
nngctagctg caccaccacc tgttcatgca ggcagagcgg ccacccctca tggaagaaga
qqaatccact qtattgggca caggetteet getggacett ggcaagcagg tgettggetg
120
gtaccaggaa gtccagcgtg tacctcagtg cgtcctcccg ataagtcctc tccaccacct
ggaacacctg gcccaacagg gtgggggctg ttgcctcaaa gggtggatac agggcggcga
240
gagtgetetg caeacagtee tecactgget caggetecat ggeteggege cgggecgegt
ccgacgettg gtcgggcggg cggggccggg cgcgccaccg cctcccttca cgcgt
355
<210> 36
<211> 118
<212> PRT
<213> Homo sapiens
<400> 36
Xaa Leu Ala Ala Pro Pro Pro Val His Ala Gly Arg Ala Ala Thr Pro
```

```
10
His Gly Arg Arg Gly Ile His Cys Ile Gly His Arg Leu Pro Ala Gly
                                25
Pro Trp Gln Ala Gly Ala Trp Leu Val Pro Gly Ser Pro Ala Cys Thr
                            40
                                                 45
Ser Val Arg Pro Pro Asp Lys Ser Ser Pro Pro Pro Gly Thr Pro Gly
                        55
Pro Thr Gly Trp Gly Leu Leu Pro Gln Arg Val Asp Thr Gly Arg Arg
                                        75
                    70
Glu Cys Ser Ala His Ser Pro Pro Leu Ala Gln Ala Pro Trp Leu Gly
                                    90
Ala Gly Pro Arg Pro Thr Leu Gly Arg Ala Gly Gly Ala Gly Arg Ala
                                105
            100
Thr Ala Ser Leu His Ala
        115
<210> 37
<211> 492
<212> DNA
<213> Homo sapiens
<400> 37
acgcgtggcc ttcgtctgcc accaggaccg actcagcccc accgggtttc cggacccgcc
gcaaccatga caagggcgat gttgtgatct gggtggattc cttctccgac atgctcgagg
gateggatet eteggeggta gteaeggtge ttgeegagge eggetatege eeaegggtee
tegeegaega egtetgetge gggttgaegt ggatcaetae eggteagete gaeggtgete
ggcgtcggct gcgcgctggt ctcgacgtgc tggcacccct gtcagacgcc agcgtcccag
togttgggct agagecgtec tgcactaccg tetggegtga tgacgcacte egecteetge
cagatgatcc gcgcgtccac cgggtagcca gaaacatgca taccgtcgcc gagatgcttg
aggeageaca gtggaeceea ceetegetag caggeeacae cetegteget cageeceatt
gtcatcccgc gg
492
<210> 38
<211> 127
<212> PRT
<213> Homo sapiens
Met Leu Glu Gly Ser Asp Leu Ser Ala Val Val Thr Val Leu Ala Glu
Ala Gly Tyr Arg Pro Arg Val Leu Ala Asp Asp Val Cys Cys Gly Leu
                                25
Thr Trp Ile Thr Thr Gly Gln Leu Asp Gly Ala Arg Arg Arg Leu Arg
                            40
Ala Gly Leu Asp Val Leu Ala Pro Leu Ser Asp Ala Ser Val Pro Val
```

```
55
Val Gly Leu Glu Pro Ser Cys Thr Thr Val Trp Arg Asp Asp Ala Leu
                    70
                                        75
Arg Leu Leu Pro Asp Asp Pro Arg Val His Arg Val Ala Arg Asn Met
His Thr Val Ala Glu Met Leu Glu Ala Ala Gln Trp Thr Pro Pro Ser
                                105
Leu Ala Gly His Thr Leu Val Ala Gln Pro His Cys His Pro Ala
                            120
<210> 39
<211> 412
<212> DNA
<213> Homo sapiens
<400> 39
aacgaaggtn ccgtacgcgc tctgaaagcc ctgcgtaaag agcgttccga tcgccgggaa
gtgatgngca ccgccaaaat gcaggtggtc gaagccgcga gttcaggcaa gattgtcttt
gaaatggaag acgtttatta cagcattgcc ggaaaacaac tggtgagcaa cttctctgcg
caagtcatgc gtggtgataa aattgcgctg attggcccga acggttgtgg taaaacgacg
ttgctgaaac tgatgttaag taagattcag gcagacagcg gccgtgttca ctgcggtact
aaactggaag ttgcgtactt cgaccagcac cgtgctgagc tggatcctga gcgtacggtg
atggataacc tggccgaagg taagcaggaa gtgatggtaa atggccgtgt an
412
<210> 40
<211> 137
<212> PRT
<213> Homo sapiens
<400> 40
Asn Glu Gly Xaa Val Arg Ala Leu Lys Ala Leu Arg Lys Glu Arg Ser
Asp Arg Arg Glu Val Met Xaa Thr Ala Lys Met Gln Val Val Glu Ala
Ala Ser Ser Gly Lys Ile Val Phe Glu Met Glu Asp Val Tyr Tyr Ser
                            40
Ile Ala Gly Lys Gln Leu Val Ser Asn Phe Ser Ala Gln Val Met Arg
                        55
                                            60
Gly Asp Lys Ile Ala Leu Ile Gly Pro Asn Gly Cys Gly Lys Thr Thr
                                        75
Leu Leu Lys Leu Met Leu Ser Lys Ile Gln Ala Asp Ser Gly Arg Val
                85
                                    90
His Cys Gly Thr Lys Leu Glu Val Ala Tyr Phe Asp Gln His Arg Ala
                                105
Glu Leu Asp Pro Glu Arg Thr Val Met Asp Asn Leu Ala Glu Gly Lys
                                                125
Gln Glu Val Met Val Asn Gly Arg Val
```

```
135
    130
<210> 41
<211> 1080
<212> DNA
<213> Homo sapiens
<400> 41
quattcaagt ggacacagge tecacgeeeg egteteaceg ataagageta caageacaac
60
tactatgacg agcgggtttc gctcgaagag cgtcttgagc gcactgtggc taaggatttc
gtcacgacgg aggtcgagcc catgtgggat gcggctgatg tcatgcggat gggtaaggat
ctetteatee ageaeggtet gacgacaaat eggaagteaa tggagtggtt taagegttae
taccccgatt tccgcgttca cgcggtgaat ttccctgggg atccgtaccc gatccatatc
gacgegacet ttgtgccgct tegtccgggg ctcatcatca acaaccegaa tegtccactg
ccqcaqqaqc agaggaagat cttcgaggcc aatgactggc agatcgttga tgctgctcag
coqqcqcacq acacqcctcc agaattgtgc tactcgtctg tgtggctatc aatgaactgc
ttggtacttg atccgaagac ggtcatctgc gaggcttcgg aagttcatca gatggagcag
atggacaagc tgggtatgaa cgtcatcccg gtcgccttcc gtgacgcgta cccattcggt
ggaggtctcc actgcgccac agctgatgta tatcgcgaag gtacctgtga ggactacttc
ccgaatcagg tcgacgaccc gaccttggtg tgagaaaacc ccgtggtcat gtcatgactg
acggateteg gtggeteggt aeggaaetta egttgteegt taeegggeeg eegggtetga
780
tatggcagta tcacgcctag caaaaaggag catgtcatgg acatggagcc gggcatcatc
aacgtcaaac aggaagttcc aggcgtcggt acgatgaacc agaaagtggg attcgtgtcc
atgettettt etgeaacggg tatggggttg gtgggtaett tegggegtet eageacteet
gtggatccca cgacgggcag taagtacatc atcggtgatt ttttggccac tggtaggatg
1020
atagtcgggg tcctgggatt tctgcttatt atcgtcatac ttggaaaatg gtctgagctc
1080
<210> 42
<211> 230
<212> PRT
<213> Homo sapiens
<400> 42
Glu Phe Lys Trp Thr Gln Ala Pro Arg Pro Arg Leu Thr Asp Lys Ser
                                    10
                 5
1
Tyr Lys His Asn Tyr Tyr Asp Glu Arg Val Ser Leu Glu Glu Arg Leu
```

```
25
Glu Arg Thr Val Ala Lys Asp Phe Val Thr Thr Glu Val Glu Pro Met
                            40
Trp Asp Ala Ala Asp Val Met Arg Met Gly Lys Asp Leu Phe Ile Gln
                        55
                                            60
His Gly Leu Thr Thr Asn Arg Lys Ser Met Glu Trp Phe Lys Arg Tyr
                    70
                                        75
Tyr Pro Asp Phe Arg Val His Ala Val Asn Phe Pro Gly Asp Pro Tyr
                85
                                    90
Pro Ile His Ile Asp Ala Thr Phe Val Pro Leu Arg Pro Gly Leu Ile
                                105
                                                    110
Ile Asn Asn Pro Asn Arg Pro Leu Pro Gln Glu Gln Arg Lys Ile Phe
                            120
Glu Ala Asn Asp Trp Gln Ile Val Asp Ala Ala Gln Pro Ala His Asp
                        135
                                            140
Thr Pro Pro Glu Leu Cys Tyr Ser Ser Val Trp Leu Ser Met Asn Cys
                    150
                                        155
Leu Val Leu Asp Pro Lys Thr Val Ile Cys Glu Ala Ser Glu Val His
                                    170
                165
Gln Met Glu Gln Met Asp Lys Leu Gly Met Asn Val Ile Pro Val Ala
            180
                                185
Phe Arg Asp Ala Tyr Pro Phe Gly Gly Gly Leu His Cys Ala Thr Ala
                                                205
                            200
Asp Val Tyr Arg Glu Gly Thr Cys Glu Asp Tyr Phe Pro Asn Gln Val
                        215
                                            220
Asp Asp Pro Thr Leu Val
225
<210> 43
<211> 358
<212> DNA
<213> Homo sapiens
<400> 43
qqqccccca catagtggac acaggtttct gggatgtcag catggagtgc caagaggtgg
gtgaccacct ggtggggaat aaggcgcttc tgggacatag aggctgcctt ccagctgcgc
120
ctggcagagc tgttgacaca acagcatggt ctgcagtgcc gggccactgc cacgcacacc
qatqtccttt aaggatggat ttgggttttc ggattcgcgt ggcctatcag cgggagtccc
agateetgaa ggaagtgeag ageecagagg ggatgatete getgagggae acagetgeet
contengent tgagagagae acaaggeagt tgecactget caccagtgee etgeacgn
<210> 44
<211> 105
<212> PRT
<213> Homo sapiens
Met Glu Cys Gln Glu Val Gly Asp His Leu Val Gly Asn Lys Ala Leu
```

```
10
1
Leu Gly His Arg Gly Cys Leu Pro Ala Ala Pro Gly Arg Ala Val Asp
            20
                                25
Thr Thr Ala Trp Ser Ala Val Pro Gly His Cys His Ala His Arg Cys
        35
                            40
                                                 45
Pro Leu Arg Met Asp Leu Gly Phe Arg Ile Arg Val Ala Tyr Gln Arg
                                            60
                        55
    50
Glu Ser Gln Ile Leu Lys Glu Val Gln Ser Pro Glu Gly Met Ile Ser
                                        75
                    70
Leu Arg Asp Thr Ala Ala Ser Leu Arg Leu Glu Arg Asp Thr Arg Gln
                                    90
Leu Pro Leu Leu Thr Ser Ala Leu His
            100
                                105
<210> 45
<211> 905
<212> DNA
<213> Homo sapiens
<400> 45
gtcgacgata aaggagtatt tgcgcagcag cagtatgatg ctctcgttga ggcgggtttc
geggeteetg gaateecaga geagtatggt ggegaeggtg eggatgegat tgegteegea
ataatcatgg aagaggtcgc tcgagtctgt gcgtcgtcgt ccaccgtcat atcgtccaat
gagettggta cegtecetet eeteaaatae ggtagegagg ageagaggaa aegttatett
totgaagttg ottogggtaa ggcactttto ggatatgege teteegagge tgatgetgga
tcagatccag ctgcacttaa gtgtcgagcc gacgaagatg gggacagttt cgtcctgaat
ggcgttaagg cttgggtcac ggaggctggc gaggccaagt acctggtgat atttgcggtt
420
actgacccag acgatecgeg ceacagaate agegegttga tggtecatge agatgacceg
ggcattaget acggggetee ggagcacaaa atggggatae gegggteagt taccagggaa
540
gtggttttca agaatacgcg tatccccaag gaacgagtaa ttggccgtcg agggcacggt
600
ctgagtgttg ctctaggtac gcttgataac tctcgtgtct cgattgctgc tcaagcagtg
ggaattgccc aaggagcttt agacattgcc acggattacg tccagaagcg caagcagttt
720
ggccagccac tgtccaattt tgagggaatc cagttcatgc tcgcagacat ggcaatgcgt
ttggaggegg egegageget gacatactet geagetgate gtagtgggeg ceagaetgae
gatgtgagtt acttcggcgc ggcggccaaa tgtttcgctt ccgacacagc gatggcagtg
900
tgcac
905
<210> 46
```

```
<211> 301
<212> PRT
<213> Homo sapiens
Val Asp Asp Lys Gly Val Phe Ala Gln Gln Gln Tyr Asp Ala Leu Val
              5
Glu Ala Gly Phe Ala Ala Pro Gly Ile Pro Glu Gln Tyr Gly Gly Asp
                             25
Gly Ala Asp Ala Ile Ala Ser Ala Ile Ile Met Glu Glu Val Ala Arg
Val Cys Ala Ser Ser Ser Thr Val Ile Ser Ser Asn Glu Leu Gly Thr
        55
Val Pro Leu Leu Lys Tyr Gly Ser Glu Glu Gln Arg Lys Arg Tyr Leu
                 70
                                   75
Ser Glu Val Ala Ser Gly Lys Ala Leu Phe Gly Tyr Ala Leu Ser Glu
Ala Asp Ala Gly Ser Asp Pro Ala Ala Leu Lys Cys Arg Ala Asp Glu
                            105
          100
Asp Gly Asp Ser Phe Val Leu Asn Gly Val Lys Ala Trp Val Thr Glu
                                           125
                        120
Ala Gly Glu Ala Lys Tyr Leu Val Ile Phe Ala Val Thr Asp Pro Asp
                    135
                                       140
Asp Pro Arg His Arg Ile Ser Ala Leu Met Val His Ala Asp Asp Pro
                                   155
                 150
Gly Ile Ser Tyr Gly Ala Pro Glu His Lys Met Gly Ile Arg Gly Ser
             165 170
Val Thr Arg Glu Val Val Phe Lys Asn Thr Arg Ile Pro Lys Glu Arg
          180
                            185
                                              190
Val Ile Gly Arg Arg Gly His Gly Leu Ser Val Ala Leu Gly Thr Leu
                        200
                                          205
Asp Asn Ser Arg Val Ser Ile Ala Ala Gln Ala Val Gly Ile Ala Gln
                                       220
                     215
Gly Ala Leu Asp Ile Ala Thr Asp Tyr Val Gln Lys Arg Lys Gln Phe
                 230
                                   235
Gly Gln Pro Leu Ser Asn Phe Glu Gly Ile Gln Phe Met Leu Ala Asp
                                250
             245
Met Ala Met Arg Leu Glu Ala Ala Arg Ala Leu Thr Tyr Ser Ala Ala
                             265
Asp Arg Ser Gly Arg Gln Thr Asp Asp Val Ser Tyr Phe Gly Ala Ala
                        280
Ala Lys Cys Phe Ala Ser Asp Thr Ala Met Ala Val Cys
                    295
  290
<210> 47
<211> 379
<212> DNA
<213> Homo sapiens
<400> 47
aagettgtag agetagteeg aageggaetg teggtaegee aagetgetaa aagatgtggg
atgcatetta eegetgegta tgeegtaget aeggaagetg ggtgeeatat eeggttaagt
```

```
cagtatgete ggaaagteeg ecagaegeag ttaagagtgg aatacetgeg eetteggetg
gegageetge etggtggtga tgetggegeg geagtaggaa ttgategteg actgegttta
gatttcgaaa aaggactcac caaatcccag ggtcgacgag aagagttcat acccgtcggc
gaagacgcca gcacgtataa cagacttatg aaagcgctgc gccaacgcca tgatgtcatc
aaatccqqaa agcttgccc
379
<210> 48
<211> 106
<212> PRT
<213> Homo sapiens
<400> 48
Met His Leu Thr Ala Ala Tyr Ala Val Ala Thr Glu Ala Gly Cys His
Ile Arg Leu Ser Gln Tyr Ala Arg Lys Val Arg Gln Thr Gln Leu Arg
                                25
Val Glu Tyr Leu Arg Leu Arg Leu Ala Ser Leu Pro Gly Gly Asp Ala
                            40
Gly Ala Ala Val Gly Ile Asp Arg Arg Leu Arg Leu Asp Phe Glu Lys
Gly Leu Thr Lys Ser Gln Gly Arg Arg Glu Glu Phe Ile Pro Val Gly
                    70
Glu Asp Ala Ser Thr Tyr Asn Arg Leu Met Lys Ala Leu Arg Gln Arg
                                    90
                85
His Asp Val Ile Lys Ser Gly Lys Leu Ala
                                105
            100
<210> 49
<211> 309
<212> DNA
<213> Homo sapiens
<400> 49
tqatcatqat gctggcatgg actattctgg tccctgttcc tctctcacct gctgaaggac
atccctctaa tttttgtgtc tccttctgta tcatcaaatt ttccctctct actgagtctc
ttgcatctcc ttggaagcat gctgtactat gtcccatcct taaagaactc cccttgtctg
cacattaccc tetgecaget ggetcatttt tetgeteece tttacaggga aactetteaa
aaaqttatct ccacctcctt ccatctcatg ttctcttgaa cctgcagtac tgggtgctcc
300
ctccttttg
309
<210> 50
<211> 101
<212> PRT
```

<213> Homo sapiens <400> 50 Met Met Leu Ala Trp Thr Ile Leu Val Pro Val Pro Leu Ser Pro Ala 10 Glu Gly His Pro Ser Asn Phe Cys Val Ser Phe Cys Ile Ile Lys Phe Ser Leu Ser Thr Glu Ser Leu Ala Ser Pro Trp Lys His Ala Val Leu 40 Cys Pro Ile Leu Lys Glu Leu Pro Leu Ser Ala His Tyr Pro Leu Pro Ala Gly Ser Phe Phe Cys Ser Pro Leu Gln Gly Asn Ser Ser Lys Ser 70 Tyr Leu His Leu Leu Pro Ser His Val Leu Leu Asn Leu Gln Tyr Trp 90 Val Leu Pro Pro Phe 100 <210> 51 <211> 512 <212> DNA <213> Homo sapiens <400> 51 agatetttga agaattgeea eactgtette etecetgett ataattteet tatteeetag qatqtqatcc ttgttcttgg ggcctcacat ggcagctgga tctctggcga ttgcatctga qttccaqaca ccaggatgga aaagaaaaga aggagggca agaggaaccc ccagatgctc cttaaqaqct actgcgtggc attcccactt gcatctcatt tgctcgatcg ctgtcactgt qccctaacqa qctgcaagga cactggggaa atgagtctgt cttgtacttc atgtgcccct caaaatcttc tgttgctgag ggagaagagg ccagccggta ttgaggaaca actagcactt tetgetteeg egteeeaggg ggaegtgggt gtgttgaate cacacegggg gtgeggaeet ctgaggctgg gctggatggg acatcaggtg ggccctctgt ttcatttatg tgacctccca tcaggtcttc tggttggatc ctgctttcta ga 512 <210> 52 <211> 125 <212> PRT <213> Homo sapiens <400> 52 Met Glu Lys Lys Arg Arg Arg Gly Lys Arg Asn Pro Gln Met Leu Leu 10 Lys Ser Tyr Cys Val Ala Phe Pro Leu Ala Ser His Leu Leu Asp Arg 20 25

Cys His Cys Ala Leu Thr Ser Cys Lys Asp Thr Gly Glu Met Ser Leu

```
40
Ser Cys Thr Ser Cys Ala Pro Gln Asn Leu Leu Leu Arg Glu Lys
                        55
Arg Pro Ala Gly Ile Glu Glu Gln Leu Ala Leu Ser Ala Ser Ala Ser
                                        75
                    70
Gln Gly Asp Val Gly Val Leu Asn Pro His Arg Gly Cys Gly Pro Leu
                85
                                    90
Arg Leu Gly Trp Met Gly His Gln Val Gly Pro Leu Phe His Leu Cys
                                105
Asp Leu Pro Ser Gly Leu Leu Val Gly Ser Cys Phe Leu
                            120
        115
<210> 53
<211> 474
<212> DNA
<213> Homo sapiens
<400> 53
accggtacac ctacgtcacc cgtaaaaacc gacgcaatac ccggatcgcc tcgtcctcaa
aaaattcgat cccgtcgtgc gtcgtcacat tgagttcaag gaggcccgct aatggccaaa
aaqtccaaqa ttgtcgccca gaagaaacgt gagaagctcg tagcccaata cgccgaaagg
cgcgccgaac tcaaggccat catgaagtgc ccaactgcct cattggacga acgcatggag
quatcageqta agetgteteg cetgeogege gatteatece cegtgeggtt acgtaacegt
qaccaaqtcg acgggcgtcc ccgcggctac gttggcaagg ccggtgtgtc ccgtatccgt
ttccqtqaga tggcccaccg cggcgaactc cccggaatcg cgaagtcaag ctggtgaagc
catggcagta ccgaagcgaa agaagtcccg ttcgaccacg cgtcataggc gggc
474
<210> 54
<211> 101
<212> PRT
<213> Homo sapiens
<400> 54
Met Ala Lys Lys Ser Lys Ile Val Ala Gln Lys Lys Arg Glu Lys Leu
Val Ala Gln Tyr Ala Glu Arg Arg Ala Glu Leu Lys Ala Ile Met Lys
Cys Pro Thr Ala Ser Leu Asp Glu Arg Met Glu Ala Ser Arg Lys Leu
                            40
Ser Arg Leu Pro Arg Asp Ser Ser Pro Val Arg Leu Arg Asn Arg Asp
                        55
Gln Val Asp Gly Arg Pro Arg Gly Tyr Val Gly Lys Ala Gly Val Ser
                                        75
Arg Ile Arg Phe Arg Glu Met Ala His Arg Gly Glu Leu Pro Gly Ile
                                    90
Ala Lys Ser Ser Trp
```

100 <210> 55 <211> 378 <212> DNA <213> Homo sapiens <400> 55 ccatggccca ggacagccgg catatcggct acgactacgg tacaccggtg gcgccacagt teggegeage caageeegea gegtgetgee aggegeaage gacaaacace ggeeegtggg tggtgttcga ccatgtgcgt tgcacccacg acacctttct gatcgacgtc tttctcaacc ageocgatge cacegogoag caggteaatg cogacaacce geactaegte gggegtttea geogeategg catgggeetg gtggatgaca agggeegttg cattacccag ggegtatege 300 gegegttgaa tgeggegege ageaccaagg egetgaacet gggaeegagt gaegeggege 360 agttatcggt gaggcgta 378 <210> 56 ' <211> 125 <212> PRT <213> Homo sapiens <400> 56 Met Ala Gln Asp Ser Arg His Ile Gly Tyr Asp Tyr Gly Thr Pro Val 10 Ala Pro Gln Phe Gly Ala Ala Lys Pro Ala Ala Cys Cys Gln Ala Gln 25 Ala Thr Asn Thr Gly Pro Trp Val Val Phe Asp His Val Arg Cys Thr 40 His Asp Thr Phe Leu Ile Asp Val Phe Leu Asn Gln Pro Asp Ala Thr 55 60 Ala Gln Gln Val Asn Ala Asp Asn Pro His Tyr Val Gly Arg Phe Ser 75 70 Arg Ile Gly Met Gly Leu Val Asp Asp Lys Gly Arg Cys Ile Thr Gln 85 90 Gly Val Ser Arg Ala Leu Asn Ala Ala Arg Ser Thr Lys Ala Leu Asn 100 105 Leu Gly Pro Ser Asp Ala Ala Gln Leu Ser Val Arg Arg 115 120 125 <210> 57

<211> 388 <212> DNA

<213> Homo sapiens

<400> 57

agacccaccc gacacagatc aggagtcgtc atgtccagaa agaagaaggt cggcatcctc 60

accgcaggcg gtgattgccc cgggctcaac gccgctatcc gcggatttgg caaggctgcc

```
atcogccage acgacatgga geteateggt atteaggacg getttettgg attggeggga
aaccqcacca totocottgg cocgegtgcc ctotcaggca tottgacggt cggcgggacc
atcctgggaa ctagccgtga caaggtcaat cacatgatta tcgacggcga ggaacgggat
atggtcccca ccaccgtcga gaattacgag aagctggggc ttgacgcttt ggtgactttg
ggtggcggtg gcaccgccaa gaacgcgt
388
<210> 58
<211> 129
<212> PRT
<213> Homo sapiens
<400> 58
Arg Pro Thr Arg His Arg Ser Gly Val Val Met Ser Arg Lys Lys
                                    10
Val Gly Ile Leu Thr Ala Gly Gly Asp Cys Pro Gly Leu Asn Ala Ala
Ile Arg Gly Phe Gly Lys Ala Ala Ile Arg Gln His Asp Met Glu Leu
Ile Gly Ile Gln Asp Gly Phe Leu Gly Leu Ala Gly Asn Arg Thr Ile
                                            60
Ser Leu Gly Pro Arg Ala Leu Ser Gly Ile Leu Thr Val Gly Gly Thr
                    70
Ile Leu Gly Thr Ser Arg Asp Lys Val Asn His Met Ile Ile Asp Gly
                85
                                    90
Glu Glu Arg Asp Met Val Pro Thr Thr Val Glu Asn Tyr Glu Lys Leu
                                105
Gly Leu Asp Ala Leu Val Thr Leu Gly Gly Gly Thr Ala Lys Asn
                            120
                                                125
        115
Ala
<210> 59
<211> 417
<212> DNA
<213> Homo sapiens
<400> 59
ggtaccatcg gagetegaca agaaatggtt gggtgaagte gtggettetg etecacecag
tgccctcatg ggtcagccca cctgaatatc ttcatgcctg tgcatttctc ctgatgttca
cgtgtgccct gtgtttttac gcatctgtga tcgtgcaccc acgcgtctca gagaggagcc
180
cgtttgggaa tccggagaat gtgcgctggc ggaagagcgt cacacactgg aagcaaacct
caqaccqcqt qqacaagacc aaqqatqaaa tggaacacga ggccttggtg gaagggaacc
300
```

```
tggcaaccga ggcaagccta gtggttctgg acacactgga gatcatcgtg cagacggtga
tgctttcaga agcccgggag agcgtcttgg gggcagtgct gaaggttgtg ctgtaca
417
<210> 60
<211> 101
<212> PRT
<213> Homo sapiens
<400> 60
Met Phe Thr Cys Ala Leu Cys Phe Tyr Ala Ser Val Ile Val His Pro
Arg Val Ser Glu Arg Ser Pro Phe Gly Asn Pro Glu Asn Val Arg Trp
                                25
Arg Lys Ser Val Thr His Trp Lys Gln Thr Ser Asp Arg Val Asp Lys
                            40
                                                 45
Thr Lys Asp Glu Met Glu His Glu Ala Leu Val Glu Gly Asn Leu Ala
Thr Glu Ala Ser Leu Val Val Leu Asp Thr Leu Glu Ile Ile Val Gln
                    70
                                         75
Thr Val Met Leu Ser Glu Ala Arg Glu Ser Val Leu Gly Ala Val Leu
                                    90
Lys Val Val Leu Tyr
            100
<210> 61
<211> 304
<212> DNA
<213> Homo sapiens
<400> 61
agatetteae ageettagae tttttteatg ggtgeettae agttttggag gteeetatee
gcacacatat ttgcaggctt ggagagagtg tgtgggggca tgtactttcg gtgggtcaag
tatgaagaag caggccttat aaacacatat tctgacctta acctgtactt cagaagagga
ccgctgactc accaaggagg cctgaaggac aaggcagcat ctctgtcttc acatgagtcc
teccetagae egggeecatg gecaggeetg accaeagage teccattgee titeetgeae
300
gcgt
304
<210> 62
<211> 92
<212> PRT
<213> Homo sapiens
<400> 62
Met Gly Ala Leu Gln Phe Trp Arg Ser Leu Ser Ala His Ile Phe Ala
                                    10
```

Gly Leu Glu Arg Val Cys Gly Gly Met Tyr Phe Arg Trp Val Lys Tyr 25 Glu Glu Ala Gly Leu Ile Asn Thr Tyr Ser Asp Leu Asn Leu Tyr Phe Arg Arg Gly Pro Leu Thr His Gln Gly Gly Leu Lys Asp Lys Ala Ala 55 Ser Leu Ser Ser His Glu Ser Ser Pro Arg Pro Gly Pro Trp Pro Gly 70 75 Leu Thr Thr Glu Leu Pro Leu Pro Phe Leu His Ala 85 90 <210> 63 <211> 577 <212> DNA <213> Homo sapiens <400> 63 egeqteaaqq qggtetacac egggacgatt aacgeetegg tgggagtatt cateacegeg ctgacggtgc tagctgggtg gctcacccta gccgggcgta tcagtgtcgg ggaactcgtc 120 accgtggtcg ggctggccca aaccctcggc cctccgctgc gagcactggg cgtcgacacc gegacgatgt tggccaccgc ccacgcctcc ggggaccgat tctgtgagtt gcgtgatagc coqqoaqoot qqoaqatoca coccqacqac qqtgcccgca ccacaccggg tgatggcccg gtggagttgc acateceggt cagggattte cagettgacg tegeeggegg caeceatgtg ggtatcatgg cgcctcaatc ggtctgtgac gccttggccg aggcgataga ccacggctcc gagaccgtct tgaatggggt teeegecagt egeeteaace etgeecaacg gegtegtetg qtqctqqtqq ctccccqctc ccccqaactq ttcqacgata ctgcccgtgc gaacatcgtg cttgacagee agacgactgt egecaggetg aatgeat 577 <210> 64 <211> 192 <212> PRT <213> Homo sapiens <400> 64 Arg Val Lys Gly Val Tyr Thr Gly Thr Ile Asn Ala Ser Val Gly Val 15 1 Phe Ile Thr Ala Leu Thr Val Leu Ala Gly Trp Leu Thr Leu Ala Gly Arg Ile Ser Val Gly Glu Leu Val Thr Val Val Gly Leu Ala Gln Thr Leu Gly Pro Pro Leu Arg Ala Leu Gly Val Asp Thr Ala Thr Met Leu 55 60 Ala Thr Ala His Ala Ser Gly Asp Arg Phe Cys Glu Leu Arg Asp Ser 70

Pro Ala Ala Trp Gln Ile His Pro Asp Asp Gly Ala Arg Thr Thr Pro 85 90 Gly Asp Gly Pro Val Glu Leu His Ile Pro Val Arg Asp Phe Gln Leu 105 Asp Val Ala Gly Gly Thr His Val Gly Ile Met Ala Pro Gln Ser Val 120 125 Cys Asp Ala Leu Ala Glu Ala Ile Asp His Gly Ser Glu Thr Val Leu 135 140 Asn Gly Val Pro Ala Ser Arg Leu Asn Pro Ala Gln Arg Arg Leu 150 155 Val Leu Val Ala Pro Arg Ser Pro Glu Leu Phe Asp Asp Thr Ala Arg 170 165 Ala Asn Ile Val Leu Asp Ser Gln Thr Thr Val Ala Arg Leu Asn Ala 185 <210> 65 <211> 339 <212> DNA <213> Homo sapiens <400> 65 gtcgaccgcg ccttgggatc gctcgaaggg gccagcctgg accaggtagc ggaagaagtc aagaaggeeg ettteaagat caeeegegee gggeaactag tgggeaeeat ggeeteegag egeettggeg taccettegg catcategae etttegettg eccetaetge egaattggga gatteggggg cecacateet tgageatatg ggattggace aagtaggeae geaeggeaea actgetgett tggetetget taacgacgee gtaaagaaag geggeatgat ggeetgeeee egegteggeg gtttgtetgg etectteate eegggetee 339 <210> 66 <211> 113 <212> PRT <213> Homo sapiens <400> 66 Val Asp Arg Ala Leu Gly Ser Leu Glu Gly Ala Ser Leu Asp Gln Val 10 Ala Glu Glu Val Lys Lys Ala Ala Phe Lys Ile Thr Arg Ala Gly Gln 25 Leu Val Gly Thr Met Ala Ser Glu Arg Leu Gly Val Pro Phe Gly Ile Ile Asp Leu Ser Leu Ala Pro Thr Ala Glu Leu Gly Asp Ser Gly Ala 55 60 His Ile Leu Glu His Met Gly Leu Asp Gln Val Gly Thr His Gly Thr 70 75 Thr Ala Ala Leu Ala Leu Leu Asn Asp Ala Val Lys Lys Gly Gly Met 90 Met Ala Cys Pro Arg Val Gly Gly Leu Ser Gly Ser Phe Ile Pro Gly 100 105

```
Ser
<210> 67
<211> 446
<212> DNA
<213> Homo sapiens
<400> 67
tgatcataaa ccacgcgtca ccgaggggat gtggcacacc tacctgcgcg tcgcagatgc
eqeaeaqqea eqqqteaqqq geqttegegg egecagetgg caeaactteg egaeeggtga
caaggggtee ttegaegeea aegagettge egtaaeteet gataetgaea eegteateea
gggagteggg ecegecetag eceteetega tteagegtgg ggaegeeaga teeaegtgga
240
gacaacaggg tgtcccagtg ccgtggtctg gaatccacgc tcctcgtcga cacatgccga
taacccgaca gcccaggcat ggcgcgattt cgtatgcgtc gagaccgggg cctgcaagga
caatgoggte attgttgccc cacacagoga cotcaccatg tocacaogga ttagogtoga
aacgttgtga tcgctgcatg gatatt
446
<210> 68
<211> 133
<212> PRT
<213> Homo sapiens
<400> 68
Met Trp His Thr Tyr Leu Arg Val Ala Asp Ala Ala Gln Ala Arg Val
                                    10
Arg Gly Val Arg Gly Ala Ser Trp His Asn Phe Ala Thr Gly Asp Lys
Gly Ser Phe Asp Ala Asn Glu Leu Ala Val Thr Pro Asp Thr Asp Thr
                            40
                                                45
Val Ile Gln Gly Val Gly Pro Ala Leu Ala Leu Leu Asp Ser Ala Trp
                        55
Gly Arg Gln Ile His Val Glu Thr Thr Gly Cys Pro Ser Ala Val Val
                                        75
Trp Asn Pro Arg Ser Ser Ser Thr His Ala Asp Asn Pro Thr Ala Gln
                                    90
Ala Trp Arg Asp Phe Val Cys Val Glu Thr Gly Ala Cys Lys Asp Asn
                                105
Ala Val Ile Val Ala Pro His Ser Asp Leu Thr Met Ser Thr Arg Ile
                            120
                                                125
Ser Val Glu Thr Leu
   130
<210> 69
<211> 552
<212> DNA
```

<213> Homo sapiens

<400> 69 nnaagggtaa ggagaaaagc aaggaccttg caaagagagc ctctgtgccg gagaggctgg ccctcaaqqa qqaqccaaaa qaaqacccca qtgqaqcaqc tgtgcccgag atgccaaaaa 120 aqtcctccaa gattgccagc ttcatcccca aaggggggaa gctcaacagt gccaagaagg 180 agenecatgg eccetteect cagtggaata ccaaaaccag gaatgaaaag catgeeeggg aaateeecaa qtqccccaqe qccttccaaq qaaqqqgage ggageeggag tgggaagetg ageteaggae teccecagea gaageeecag etggaeggea gacaetecag tteetettee agectggegt ceteagaagg aaaaggeeca ggagggaeca ceetgaacca cageateage agccagactg tcagtgggtc tgtcgggacc acccagacca caggaagcaa tnnaccgtca qtqttcaqct acctcaqccc caqcaqcaat acaaccatcc caacactgcc acggttgcac 540 ctttcctgta ca 552 <210> 70 <211> 184 <212> PRT <213> Homo sapiens <400> 70 Xaa Arg Val Arg Arg Lys Ala Arg Thr Leu Gln Arg Glu Pro Leu Cys 15 1 10 Arg Arg Gly Trp Pro Ser Arg Arg Ser Gln Lys Lys Thr Pro Val Glu Gln Leu Cys Pro Arg Cys Gln Lys Ser Pro Pro Arg Leu Pro Ala Ser 40 45 Ser Pro Lys Gly Gly Ser Ser Thr Val Pro Arg Arg Ser Xaa Met Ala 55 60 Pro Ser Leu Ser Gly Ile Pro Lys Pro Gly Met Lys Ser Met Pro Gly 70 75 Lys Ser Pro Ser Ala Pro Ala Pro Ser Lys Glu Gly Glu Arg Ser Arg 90 Ser Gly Lys Leu Ser Ser Gly Leu Pro Gln Gln Lys Pro Gln Leu Asp 105 Gly Arg His Ser Ser Ser Ser Ser Leu Ala Ser Ser Glu Gly Lys 120 125 Gly Pro Gly Gly Thr Thr Leu Asn His Ser Ile Ser Ser Gln Thr Val 135 Ser Gly Ser Val Gly Thr Thr Gln Thr Thr Gly Ser Asn Xaa Pro Ser 150 155 Val Phe Ser Tyr Leu Ser Pro Ser Ser Asn Thr Thr Ile Pro Thr Leu 175 165 170 Pro Arg Leu His Leu Ser Cys Thr 180

<210> 71

```
<211> 316
<212> DNA
<213> Homo sapiens
<400> 71
cgcgttgaaa tggcgttcga acttaaacgt ttacatattg actccgtgcc attaaacatt
ttgaatcctg ttaaagggac tccatttgaa agcaacgaag ctttacgtcc tttaaatatc
ttacqtacct tcqccqtatt ccqtttcatc ttgccaaacg cattgatacg aactgcaggt
ggccgcgaag taaatctacg agacttacaa gcttatgctc taaaaggtgg cctaaacggt
240
atcatggttg gtggctactt aactactggc ggtcgttcac ctcaagacga tctccaaatg
attcaagact tggagt
316
<210> 72
<211> 105
<212> PRT
<213> Homo sapiens
Arg Val Glu Met Ala Phe Glu Leu Lys Arg Leu His Ile Asp Ser Val
                 5
1
Pro Leu Asn Ile Leu Asn Pro Val Lys Gly Thr Pro Phe Glu Ser Asn
                                25
            20
Glu Ala Leu Arg Pro Leu Asn Ile Leu Arg Thr Phe Ala Val Phe Arg
                                                45
                            40
Phe Ile Leu Pro Asn Ala Leu Ile Arg Thr Ala Gly Gly Arg Glu Val
Asn Leu Arg Asp Leu Gln Ala Tyr Ala Leu Lys Gly Gly Leu Asn Gly
Ile Met Val Gly Gly Tyr Leu Thr Thr Gly Gly Arg Ser Pro Gln Asp
                                    90
                85
Asp Leu Gln Met Ile Gln Asp Leu Glu
           100
<210> 73
<211> 384
<212> DNA
<213> Homo sapiens
nntaccqqca agatcctggc cgaaggtgac gtcgaggttt ctgaggctat cgactttgct
gettggtatg tegacegage egaggagete gagggegteg aeggtgeeca gtttgtgeeg
ccacqagtga ccgtcgtcac cccgccgtgg aacttcgccc tgtctattac cgccggatcc
```

```
accettgeeg etetggeege eggategtea gtaetaetea ageeegetee acaggeeege
240
cactgtgctg ccgtcatctc tgaatgcctg tgggaggctg ggatcccgcg ggacgttctg
careteqteq atqttqaqqa aaatgagget ggtaaacace tggtgageca ceeegaggte
gatcgggtca tcctcacggg aggt
384
<210> 74
<211> 128
<212> PRT
<213> Homo sapiens
<400> 74
Xaa Thr Gly Lys Ile Leu Ala Glu Gly Asp Val Glu Val Ser Glu Ala
                                    10
                                                        15
Ile Asp Phe Ala Ala Trp Tyr Val Asp Arg Ala Glu Glu Leu Glu Gly
                                25
Val Asp Gly Ala Gln Phe Val Pro Pro Arg Val Thr Val Val Thr Pro
                            40
Pro Trp Asn Phe Ala Leu Ser Ile Thr Ala Gly Ser Thr Leu Ala Ala
Leu Ala Ala Gly Ser Ser Val Leu Leu Lys Pro Ala Pro Gln Ala Arg
His Cys Ala Ala Val Ile Ser Glu Cys Leu Trp Glu Ala Gly Ile Pro
                85
Arg Asp Val Leu Gln Leu Val Asp Val Glu Glu Asn Glu Ala Gly Lys
            100
                                105
                                                    110
His Leu Val Ser His Pro Glu Val Asp Arg Val Ile Leu Thr Gly Gly
                                                125
        115
                            120
<210> 75
<211> 405
<212> DNA
<213> Homo sapiens
<400> 75
qaattcgtct cggaatacac gctggaaaat tcggccgaga tgtccggggt gcgctcanac
cgcattgagg cgctggccga gctctatgcc gatcccaaga ccagggtggt gagcttctgg
120
accatgggct tcaaccagca cacccgcggc gtctggtgca acaatctcgt ctacaacatc
cacctgctga ccggaaaaat ctcgacgccc ggcaacagcc cgttctcgct gaccgggcag
ccatcggcct gcggcacggc gcgcgaggtc ggtaccttct cgcatcgcct gcccgccgac
atggtggtca ccagcaaggc gcaccgcgac atcgccgaga agatctggca gctgccggaa
ggaccagtec ecgacaagee eggetaceae geegtgetge agage
405
<210> 76
```

```
<211> 135
<212> PRT
<213> Homo sapiens
<400> 76
Glu Phe Val Ser Glu Tyr Thr Leu Glu Asn Ser Ala Glu Met Ser Gly
                                    10
Val Arg Ser Xaa Arg Ile Glu Ala Leu Ala Glu Leu Tyr Ala Asp Pro
            20
                                25
Lys Thr Arg Val Val Ser Phe Trp Thr Met Gly Phe Asn Gln His Thr
                            40
Arg Gly Val Trp Cys Asn Asn Leu Val Tyr Asn Ile His Leu Leu Thr
                        55
Gly Lys Ile Ser Thr Pro Gly Asn Ser Pro Phe Ser Leu Thr Gly Gln
                                        75
                    70
Pro Ser Ala Cys Gly Thr Ala Arg Glu Val Gly Thr Phe Ser His Arg
Leu Pro Ala Asp Met Val Val Thr Ser Lys Ala His Arg Asp Ile Ala
            100
Glu Lys Ile Trp Gln Leu Pro Glu Gly Pro Val Pro Asp Lys Pro Gly
                            120
        115
Tyr His Ala Val Leu Gln Ser
    130
<210> 77
<211> 5816
<212> DNA
<213> Homo sapiens
<400> 77
gageggegee etgetetgge egttgeteee geteetgete etgetgetgt eggegeggga
eggegtgege geegegeage eteaggeece gggttacttg attgeagete cetetgtttt
tegegeggge gtggaggaag teateagegt gaccatettt aacteteeaa gggaagteae
ggtccaggct cagctggtgg cccagggtga gccggtggtg cagagccagg gagccatcct
ggataaaggg acaatcaaac tcaaggtgcc cacgggcctc cggggccaag cgcttctgaa
agtgtggggc cgcggctggc aggcggagga ggggcccctc tttcacaacc agacctcggt
gaccgtggac ggccggggcg cttctgtatt catccagacg gacaagcctg tgtacagacc
ccagcaccga gtgctcataa gcatcttcac cgtctctcca aatctgaggc ctgtcaacga
gaagetggaa geetacatee tggaceeeeg aggetetegg atgatagagt ggagacaett
aaageegtte tgetgeggea teaceaacat gagetteece ttgteegace ageetgtgtt
gggagaatgg ttcatttttg ttgaaatgca aggccacgcg tacaacaagt cttttgaagt
tcagaagtat gtgttgccca agtttgagct tctgattgac ccgccccggt atatccaaga
720
```

cctggacgcc 780	tgtgagacag	gcactgtgcg	ggccaggtat	acctttggga	aacctgtggc
	atgatcaaca	tgactgttaa	tggtgtaggg	tactacagec	acgaggtggg
acgccctgtc 900	ctcagaacaa	ccaagatcct	cggctcccgg	gacttcgaca	tctgcgtgag
960			cttccggggc		
1020			cgcgttcgat		
1080			cacgaggaag		•
1140			cgatggcagc		
1200			taacatctac		
1260			catececaeg		
1320			gcccgtgggg		
1380			ccagtgctac		
1440			tttttctgtg		
1500			gggcaatatt		
1560			ggcggcccct		
1620			accageceea		
tgtgacctct 1680	cttcatctgg	ccgtgacccc	cagcatggtc	ccccttggtc	gcctgctggt
cttctacgtc 1740	agggagaatg	gagaaggggt	cgccgacagc	cttcagtttg	cagtcgagac
cttcttcgaa 1800	aaccaggttt	cagtgacgta	ttcagcaaat	gagacccaac	ctggggaggt
tgtcgacctg	cggatcaggg	ctgcaagggg	cagctgtgtg	tgcgtcgccg	cagttgataa
gagtgtctac 1920	ctgctcaggt	ctgggttccg	getgaeteet	gcccaggttt	tccaggaact
ggaagattat 1980	gatgtttctg	attcctttgg	cgtgtccagg	gaggatggtc	ctttttggtg
	acggcacaac	gacgccggcg	ctcctctgtc	ttcccgtggc	cttggggcat
	tctgggtttg	ccttcaccga	aacgggactg	gtggtgatga	ccgaccgagt
	caccggcagg	acggtggcct	ctacaccgat	gaggctgtcc	ccgctttcca
	gggagcctgg	tggcagtggc	tccttccagg	caccccccca	gaacagagaa
	actttcttcc	ccgaaacatg	gatttggcat	tgtctcaaca	tcagtgaccc
	gggacactca	gtgtgaaggt	cccggactcc	atcaccagct	gggtgggtga

2400				gagccctccc	
cttcaagccc 2460	ttcttcgtgg	acttcatgct	cccgctctc	atcatccgtg	gggagcaggt
caagatcccg 2520	ctcagtgtct	acaactacat	gggcacctgc	gctgaggtgt	acatgaagct
ctcggttccc 2580	aagggcatcc	agtttgttgg	gcatcctggc	aaacgccatg	tgaccaagaa
gatgtgtgtg	gcccccgggg	aggctgagcc	catctgggtc	gttctgtcct	tcagcgacct
	aacatcacgg	ccaaagccct	tgcttacgga	gacacaaatt	gctgccggga
	agcaaacacc	ctgaggagaa	tcacgccgac	aggagggtcc	ccatcggggt
	aggegeagtg	tgatggttga	ggcggaagga	gteccecggg	cgtacaccta
cagcgcattc 2880	ttetgtecca	gtgagagagt	ccacatctcc	acccccaaca	agtatgagtt
ccagtatgtg 2940	cagcggccac	tgcgcctcac	ccgctttgat	gtggctgtgc	gagctcacaa
tgatgcccgt	gtggccttgt	cttctgggcc	ccaggacaca	gcaggcatga	tcgagatcgt
	catcagaaca	ccaggtcatg	gatetecace	agcaagatgg	gagageeegt
	cacacggcca	agatcctctc	ctgggatgaa	ttcagaacat	tctggatcag
	ggccttatcc	aggttggcca	tggtccagag	ccatccaatg	agtctgtcat
	accctcccga	ggccaccaga	ggtccagttc	attggctttt	ccaccggctg
	ggtgaattcc	gaatctggag	gaagatggag	gtggacgaga	gctacagcga
	tgggggtccc	acacggcgcc	atccctgggt	ctgagcgagc	caccgcctcc
	acgtcatggg	gccaaccctg	aaccacctca	acaacctcct	geggetgeeg
	gagagcagaa	catgatccac	tttgcaccca	acgtctttgt	cttgaagtat
	cccagcagct	cagccctgag	gtggagagag	agaccaccga	ctacctagta
	agcgccagct	gacctacaag	cgccaggatg	gctcctacag	cgcgtttggg
-	catcggggag	catgtggctc	acagcctttg	tcctgaagtc	cttcgcacag
	ttatcttcgt	ggacccccgg	gagctggctg	ccgccaagag	ctggatcatc
	aggccgatgg	ctccttcctg	gccgtgggca	gggtcctgaa	caaggacatc
3780 cagggtggga	tccacggcat	tgtecegetg	acagcctacg	tggtggttgc	tctcctggaa
3840 acaggcacag	cctcagagga	ggagagaggc	tccactgaca	aagcgaggca	cttcctggag
	ccctggccat	ggaccettat	agetgtgccc	tgactaccta	cgcgctgacc
3960					

ctgctccgca 4020	gcccggcagc	ccctgaggca	ctgcgcaagc	teegtageet	ggccatcatg
cgagatgggg 4080	tcacccactg	gagcctgtca	aattcctggg	acgtggacaa	gggcacattc
ttgagcttca 4140	gtgacagggt	ctctcagtca	gtggtctcgg	ccgaggtgga	aatgacagcc
tacgcccttc 4200	tgacctacac	tctgctgggt	gacgtggctg	ccgccctgcc	tgtggtgaag
	agcagcgaaa	tgcacttggg	ggetteteet	ccactcagga	cacctgcgtg
	ccttggctga	atatgccatc	ttgtcctatg	ctggaggcat	caacctcact
gtctccctgg 4380	cctccaccaa	cctggactac	caggaaacct	tcgagctgca	caggaccaac
	tgcagacagc	agcgatcccc	agectececa	cggggctgtt	tgtgagtgcc
aagggggacg 4500	getgetgeet	gatgcagatt	gatgtcacct	acaatgtgcc	tgacccggtg
gccaagccag 4560	ctttccagct	gctcgtaagc	ctccaggagc	ctgaggccca	gggacgcccg
cccccatgc 4620	ctgcctccgc	agctgagggt	tcccgaggag	actggccccc	agetgaegat
4680			gaatacaagg		
4740			getgteetgg		
4800			ctgctccttg		
4860			ttctactttg		
4920			gagtgcgtgg		
4980			gccttcgagg		
5040			tgcgccggac		
5100			cccggcgagt		
5160			tgcgaccacg		
5220			gccagcgcct		
5280			cctcccagct		
5340			gatgacctgg		
5400					agaccctgag
5460			gccgggcctc		
5520					gactccggcc
cctcagagac 5580	atagtggccg	ggtggtgggg	geceacagge	cagggcttct	gagccctgtc

ttcgtctaca gcccagcctt tcagagtggt ggggaggagg gtttatggat gtcaaacacc tgcaccttga gataatccta caaccacatg cagttgtggg accgcagttt ggtcctgggg 5700 accattcata eccacacace cagettgtge etgtggttaa catetcagaa aactetggta aatgatcact ccaggatatt gacacgaata cacgttactg atcttactca catgtt 5816 <210> 78 <211> 799 <212> PRT <213> Homo sapiens <400> 78 Ile Pro Asn Leu Glu Glu Asp Gly Gly Gly Arg Glu Leu Gln Arg Gly Leu His Leu Gly Val Pro His Gly Ala Ile Pro Gly Ser Glu Arg Ala 25 Thr Ala Ser Ile Ile Gly Asp Val Met Gly Pro Thr Leu Asn His Leu 40 Asn Asn Leu Leu Arg Leu Pro Phe Gly Cys Gly Glu Gln Asn Met Ile 55 His Phe Ala Pro Asn Val Phe Val Leu Lys Tyr Leu Gln Lys Thr Gln 75 70 Gln Leu Ser Pro Glu Val Glu Arg Glu Thr Thr Asp Tyr Leu Val Gln 90 85 Gly Tyr Gln Arg Gln Leu Thr Tyr Lys Arg Gln Asp Gly Ser Tyr Ser 110 105 Ala Phe Gly Glu Arg Asp Ala Ser Gly Ser Met Trp Leu Thr Ala Phe 120 125 Val Leu Lys Ser Phe Ala Gln Ala Arg Ser Phe Ile Phe Val Asp Pro 140 135 Arg Glu Leu Ala Ala Lys Ser Trp Ile Ile Gln Gln Gln Ala 155 150 Asp Gly Ser Phe Leu Ala Val Gly Arg Val Leu Asn Lys Asp Ile Gln 170 Gly Gly Ile His Gly Ile Val Pro Leu Thr Ala Tyr Val Val Val Ala 185 180 Leu Leu Glu Thr Gly Thr Ala Ser Glu Glu Glu Arg Gly Ser Thr Asp 200 Lys Ala Arg His Phe Leu Glu Ser Ala Ala Pro Leu Ala Met Asp Pro 215 220 Tyr Ser Cys Ala Leu Thr Thr Tyr Ala Leu Thr Leu Leu Arg Ser Pro 235 230 Ala Ala Pro Glu Ala Leu Arg Lys Leu Arg Ser Leu Ala Ile Met Arg 250 245 Asp Gly Val Thr His Trp Ser Leu Ser Asn Ser Trp Asp Val Asp Lys 265 260 Gly Thr Phe Leu Ser Phe Ser Asp Arg Val Ser Gln Ser Val Val Ser 285 280 275 Ala Glu Val Glu Met Thr Ala Tyr Ala Leu Leu Thr Tyr Thr Leu Leu 300 295 Gly Asp Val Ala Ala Ala Leu Pro Val Val Lys Trp Leu Ser Gln Gln

305					310					315					320
	Asn	Ala	Leu	Gly 325		Phe	Ser	Ser	Thr		Asp	Thr	Cys	Val 335	Ala
Leu	Gln	Ala	Leu 340	Ala	Glu	Tyr	Ala	Ile 345	Leu	Ser	Tyr	Ala	Gly 350	Gly	Ile
Asn	Leu	Thr 355	Val	Ser	Leu	Ala	Ser 360	Thr	Asn	Leu	Asp	Tyr 365	Gln	Glu	Thr
	370		His	_		375		-			380				
385			Pro		390					395					400
_			Gln	405	_			_	410					415	
_			Phe 420					425					430		
_		435	Pro				440					445			
_	450		Pro Lys			455					460				
465			Asn		470					475					480
_			Ile	485					490					495	
			500 Arg			•		505					510		
_		515	Pro				520					525			
	530		Val			535					540				
545			Glu		550					555					560
-	-	-	Pro	56 5					570					575	
			580 Arg			_		585	_				590		
		5 95	Ala				600					605			
	610		His			615					620				
625			Val		630					635					640
_			Ala	645					650					655	
_			660 Leu					665					670		
		675	Ala				680					685			
Ala	690 Gly	Leu	Glu	Val	Glu	695 Asp	Ser	Asp	Pro	Glu	700 Pro	Glu	Gly	Glu	Ala
705	_		Val		710					715					720
	_		Ser	725					730					735	
_												_		-	

```
740
                                745
Thr Pro Ala Pro Gln Arg His Ser Gly Arg Val Val Gly Ala His Arg
                            760
Pro Gly Leu Leu Ser Pro Val Phe Val Tyr Ser Pro Ala Phe Gln Ser
                        775
                                            780
Gly Gly Glu Glu Gly Leu Trp Met Ser Asn Thr Cys Thr Leu Arg
                    790
785
<210> 79
<211> 346
<212> DNA
<213> Homo sapiens
<400> 79
acgcgttgcc gtcggggaag tggtagatgt aaggctcttc catttctgcg gcggcttctt
cgctgatagg cgcgaggttc atgccgccaa agcccgggaa acccacgctg taggcatccg
ccgatgcaac ggcatcaata ctcaaggcca tgaagtcgtt ggtgcggctc tggaagtact
tcaccqcqcc ttccgacaag cccacqtcct tgagcaggaa gtcgcggtag ctggtggccg
ccagatactc ggctttttct tcggcggact tgcccggcag gtaatccttg ggcgcgacgt
qcatqqcqat caatqcccgg ctgtccgctt ccggncnnnc cnncnn
346
<210> 80
<211> 101
<212> PRT
<213> Homo sapiens
<400> 80
Met His Val Ala Pro Lys Asp Tyr Leu Pro Gly Lys Ser Ala Glu Glu
Lys Ala Glu Tyr Leu Ala Ala Thr Ser Tyr Arg Asp Phe Leu Leu Lys
                                25
            20
Asp Val Gly Leu Ser Glu Gly Ala Val Lys Tyr Phe Gln Ser Arg Thr
                            40
Asn Asp Phe Met Ala Leu Ser Ile Asp Ala Val Ala Ser Ala Asp Ala
                        55
Tyr Ser Val Gly Phe Pro Gly Phe Gly Gly Met Asn Leu Ala Pro Ile
                                        75
                    70
Ser Glu Glu Ala Ala Glu Met Glu Glu Pro Tyr Ile Tyr His Phe
                                    90
Pro Asp Gly Asn Ala
            100
<210> 81
<211> 429
<212> DNA
<213> Homo sapiens
<400> 81
```

```
nngcttttct cnccctcgca cctgatctgc tgctgacgtg cgggagggtc attaaaagat
60
tgacttaaat ttgataccag aaatgtcagt gtggctttgc ttgaacagcc gcgtgcgggg
120
cetetatgag tggaatecag teteatggee ecceecatgg eteetgttae eetggaggag
getactecat gaggetgget eeggtggeae tgeetegggg etgeatgtae atgtgtgtge
240
atgnntgtgt gcatgtgcgt gtgcacgtgt nngtgtgtgc ccgtgtgcat gtgcccgtgt
300
gegtgtetgt getgtgtgtg tgegtgeatg tgtgegtgte tgtgegtgne tgtgtgtgtg
cgtgcatgtg tgtgcacgtg tgtgcatgtg cacgtgtgcg cacccgtgtg catgtccgtg
420
tgcacgcgt
429
<210> 82
<211> 79
<212> PRT
<213> Homo sapiens
<400> 82
Gly Trp Leu Arg Trp His Cys Leu Gly Ala Ala Cys Thr Cys Val Cys
Met Xaa Val Cys Met Cys Val Cys Thr Cys Xaa Cys Val Pro Val Cys
                                 25
            20
Met Cys Pro Cys Ala Cys Leu Cys Cys Val Cys Ala Cys Met Cys Ala
                            40
Cys Leu Cys Val Xaa Val Cys Val Arg Ala Cys Val Cys Thr Cys Val
                        55
                                             60
    50
His Val His Val Cys Ala Pro Val Cys Met Ser Val Cys Thr Arg
                                         75
                    70
65
<210> 83
<211> 411
<212> DNA
<213> Homo sapiens
<400> 83
gtgcacttct tcggacgagg gtccttcggc gcccgcgata acgagagggt gcggagattt
cgtgcgctct ttaaaaagcg caaaggcttt tataagttcg atgtgctttt tttcaggccc
ggaaagacgc gatccgtaga taaaataggg ccgctttatc gcaaatggct ttatggttac
180
cgtttccgca tcgatttgca gctgcggaaa aaaaggcctg tggtcgatgc cgttgtggat
gatttegatt ttegetttgt egatgecaag eeetateaaa tegtttegea aaaacteget
cgcggcgacg atcttttgca cctttgtaag cccccgcttt atctgcatgc gcatcatgcg
gtcgctcttt tctttgagca tgacggaaac gatggaattg acgacggcga c
411
```

```
<210> 84
<211> 127
<212> PRT
<213> Homo sapiens
<400> 84
Met Leu Lys Glu Lys Ser Asp Arg Met Met Arg Met Gln Ile Lys Arg
                                   10
                5
Gly Leu Thr Lys Val Gln Lys Ile Val Ala Ala Ser Glu Phe Leu Arg
                                25
Asn Asp Leu Ile Gly Leu Gly Ile Asp Lys Ala Lys Ile Glu Ile Ile
                           40
His Asn Gly Ile Asp His Arg Pro Phe Phe Pro Gln Leu Gln Ile Asp
Ala Glu Thr Val Thr Ile Lys Pro Phe Ala Ile Lys Arg Pro Tyr Phe
                                        75
                    70
Ile Tyr Gly Ser Arg Leu Ser Gly Pro Glu Lys Lys His Ile Glu Leu
                                    90
                85
Ile Lys Ala Phe Ala Leu Phe Lys Glu Arg Thr Lys Ser Pro His Pro
                                105
           100
Leu Val Ile Ala Gly Ala Glu Gly Pro Ser Ser Glu Glu Val His
                            120
<210> 85
<211> 333
<212> DNA
<213> Homo sapiens
<400> 85
nnacgcgtgc cgcgcaaggg aaccatgttc gtgtcggtgc gcgaaaccga caaggcgcgc
atcetegacg eggtgaaact getgagtteg eteggettea aggtgatege eaceteggge
acccagcgtt tcctggtgga gaacggagta ccggcggaaa agatcaacaa ggtgctggaa
ggccgcccgc acattgttga cgcaattacc aacggcgagg tgcaactcgt tttcaatacc
accgaggggc cacaggcgct ggctgacagc cgctcgttgc gacgcgctgc cctcttgcat
aaagtgccat attacaccac tctttcaggt gca
333
<210> 86
<211> 111
<212> PRT
<213> Homo sapiens
<400> 86
Xaa Arg Val Pro Arg Lys Gly Thr Met Phe Val Ser Val Arg Glu Thr
                                    10
                 5
Asp Lys Ala Arg Ile Leu Asp Ala Val Lys Leu Leu Ser Ser Leu Gly
                                25
Phe Lys Val Ile Ala Thr Ser Gly Thr Gln Arg Phe Leu Val Glu Asn
```

ATTORNEY DOCKET NO.: 15966-543

40 Gly Val Pro Ala Glu Lys Ile Asn Lys Val Leu Glu Gly Arg Pro His Ile Val Asp Ala Ile Thr Asn Gly Glu Val Gln Leu Val Phe Asn Thr Thr Glu Gly Pro Gln Ala Leu Ala Asp Ser Arg Ser Leu Arg Arg Ala 85 90 Ala Leu Leu His Lys Val Pro Tyr Tyr Thr Thr Leu Ser Gly Ala 105 <210> 87 <211> 355 <212> DNA <213> Homo sapiens acgcgtgagg aaatggggge cgcaggcctg gatcgcaagg tatggcagtg cccggtcgtc ctcctgagcg atgttcattc ggtaggggta cagggtgacg ggcgtactta tggttctccc attgtgcttc gcccggtgac gagtgaggac gccatgactg cggactgggc acgtatccca tatgacgtac tggaaaagat ctcgactcgc attacgaatg cgtgtccgca aatcaaccgg gtggtactcg atatcacatc taaaccgccg gccaccatcg agtgggaatg agccccgtct caccgtgaac atgacatggc ccgcaccttt cttggggcgg gccatgccgt gttag 355 <210> 88 <211> 96 <212> PRT <213> Homo sapiens <400> 88 Thr Arg Glu Glu Met Gly Ala Ala Gly Leu Asp Arg Lys Val Trp Gln 10 Cys Pro Val Val Leu Leu Ser Asp Val His Ser Val Gly Val Gln Gly 25 Asp Gly Arg Thr Tyr Gly Ser Pro Ile Val Leu Arg Pro Val Thr Ser 40 Glu Asp Ala Met Thr Ala Asp Trp Ala Arg Ile Pro Tyr Asp Val Leu Glu Lys Ile Ser Thr Arg Ile Thr Asn Ala Cys Pro Gln Ile Asn Arg 75 70 Val Val Leu Asp Ile Thr Ser Lys Pro Pro Ala Thr Ile Glu Trp Glu 85 90 <210> 89

<211> 351

(211) 331

<212> DNA

<213> Homo sapiens

<400> 89

ATTORNEY DOCKET NO.: 15966-543

nacgcgtcaa caccaggcta cggtgggtat gatcatgata agggctggga cccgcaggag agggcaagag ttatataata gacgggtgta tgggtttaag gtatacatgg tgatgagaca qqtgattgga taaggccaga gggttgggcg gggttctgcc cctgctgaag cctggtgggg cecaqqtctq tqatctggga ccggaacaac acatctgctc tgggcctgct ggatgtggcg caagccctgg aacagaacca cagcctcaag tccatgccgc tgccactgaa tgacgtaacc caggeteate geageeggee agaacteaca actegagegg tecateagat e 351 <210> 90 <211> 61 <212> PRT <213> Homo sapiens <400> 90 Ser Leu Val Gly Pro Arg Ser Val Ile Trp Asp Arg Asn Asn Thr Ser 10 Ala Leu Gly Leu Leu Asp Val Ala Gln Ala Leu Glu Gln Asn His Ser 25 Leu Lys Ser Met Pro Leu Pro Leu Asn Asp Val Thr Gln Ala His Arg Ser Arg Pro Glu Leu Thr Thr Arg Ala Val His Gln Ile 50 <210> 91 <211> 327 <212> DNA <213> Homo sapiens <400> 91 ngqtgggtcc tctcgcagga caggtaatcc tgagacctac gcagctccct ggagctctgc acgtcgaagc ccagcagggc ctcctgcagg tccctggggc agccagcaca cacaaagtcc cggaagggc tgtagactcc ctgccagcgg ctttcccggg gaaggcaccc acgcccagct geotettgea ggtactgete gggtetggtg ggagggeage egtgteeage acaccetgtg tgtgcagtcc tctccctgcc ccactgccga acgagccctc cacggtgaag ccattgggga 300 acgtgacctt gcccttcccc atgaggt 327 <210> 92 <211> 107 <212> PRT <213> Homo sapiens Met Gly Lys Gly Lys Val Thr Phe Pro Asn Gly Phe Thr Val Glu Gly

ATTORNEY DOCKET NO.: 15966-543

10 Ser Phe Gly Ser Gly Ala Gly Arg Gly Leu His Thr Gln Gly Val Leu 25 Asp Thr Ala Ala Leu Pro Pro Asp Pro Ser Ser Thr Cys Lys Arg Gln 45 40 Leu Gly Val Gly Ala Phe Pro Gly Lys Ala Ala Gly Arg Glu Ser Thr Ala Pro Ser Gly Thr Leu Cys Val Leu Ala Pro Gly Thr Cys Arg Arg Pro Cys Trp Ala Ser Thr Cys Arg Ala Pro Gly Ser Cys Val Gly 85 90 Leu Arg Ile Thr Cys Pro Ala Arg Gly Pro Thr <210> 93 <211> 394 <212> DNA <213> Homo sapiens <400> 93 neegegtacg acaageagta cetegagggg egttacggtg eggacecata cetgageaac atgctcgaat gggacggcgg acatgagcaa taggccgcca aagcacagcg agaggaaggc ggcgcgctct gcggcctcgc ggatgaagga ggggtggtcg cgcaagagac gccccgcgcc ttttgtaccg ataaatcccg ggcgccacct gatcgtgacc gagggtacga aaacggaacc getetattte gaggetatea ggttgegtgt caacaacegt tateaeggee agtgggtgae aatggaagtt gtcgttaccg gcaagcatac caggggactt ctcgatcgtg cagtcactct ggcggaagaa agtgccacag gattcactca cgta 394 <210> 94 <211> 109 <212> PRT <213> Homo sapiens <400> 94 Met Leu Ala Gly Asn Asp Asn Phe His Cys His Pro Leu Ala Val Ile 10 Thr Val Val Asp Thr Gln Pro Asp Ser Leu Glu Ile Glu Arg Phe Arg 25 Phe Arg Thr Leu Gly His Asp Gln Val Ala Pro Gly Ile Tyr Arg Tyr 45 40 Lys Arg Arg Gly Ala Ser Leu Ala Arg Pro Pro Leu Leu His Pro Arg Gly Arg Arg Ala Arg Arg Leu Pro Leu Ala Val Leu Trp Arg Pro Ile 75 70

90

Ala His Val Arg Arg Pro Ile Arg Ala Cys Cys Ser Gly Met Gly Pro

His Arg Asn Ala Pro Arg Gly Thr Ala Cys Arg Thr Arg

85

105

100 <210> 95 <211> 531 <212> DNA <213> Homo sapiens <400> 95 ggtacctctt ccaagtacct tctaaatgaa acactcaaga gagtgctact caggaaactt 60 tgcttggatc ctaaaatgga ctggtcttgg gtgtgtaacc ccggtgaagt tatagcctcc 120 ccaaattgag gtgacagaag gaagacaaga ggtgtaagct ggagagggaa gggaagaaat 180 cagtggettt ggccageete tgtgccaeee agtacgaeag aggagtggga aetggeeete tggggctctg cttggccata ggcactgcac attgtgccac ctgctcatca cctcctctag teteacactg ageateggag tacetgttgt geagacagga aaaetgagga getetgagag getgageatg gageteacce catgecatag ggtgtgggaa gagggeacag gaggeeteat ccatggggga aagggttgag gatggacatg ggtggggaga gggcatagac atcccttcct aatctctgtt cccaccacat ttcataggag atgagttagg agatgacagc t 531 <210> 96 <211> 124 <212> PRT <213> Homo sapiens <400> 96 Met Arg Pro Pro Val Pro Ser Ser His Thr Leu Trp His Gly Val Ser 10 5 Ser Met Leu Ser Leu Ser Glu Leu Leu Ser Phe Pro Val Cys Thr Thr 20 25 Gly Thr Pro Met Leu Ser Val Arg Leu Glu Glu Val Met Ser Arg Trp His Asn Val Gln Cys Leu Trp Pro Ser Arg Ala Pro Glu Gly Gln Phe 55 Pro Leu Leu Cys Arg Thr Gly Trp His Arg Gly Trp Pro Lys Pro Leu 75 70 Ile Ser Ser Leu Pro Ser Pro Ala Tyr Thr Ser Cys Leu Pro Ser Val 90 85 Thr Ser Ile Trp Gly Gly Tyr Asn Phe Thr Gly Val Thr His Pro Arg 105 100 Pro Val His Phe Arg Ile Gln Ala Lys Phe Pro Glu 120 <210> 97 <211> 405 <212> DNA <213> Homo sapiens

<400> 97

```
ggtcgggcca gtcgaacttc attcccgctt cgagggtctt gctgcggatg ggccgtacgc
tcgcggtgcc ttgcgcgcgg gctggtaggt ggagaagccg cgcgagtacg cgccgtagag
cgacatcgtg tctgagacgt cgaagctcag gcccagcttt ggcgtccagg cgcgctcggt
cggtccgccc tcttgcggca attgattcag cgcaatcccg gccatcacat gccagcgctt
gtccagggtc atgaaatcct gggcataggc gcgcgaggag cgcagcggcg aattggacag
300
gcgctcgata ttgggcgtga tgtccgaaga cgggaacggg acccgggggg agaagacgtt
gcccgggaaa agatcccccg acgccatcgt ggtgtcgacc gagat
405
<210> 98
<211> 122
<212> PRT
<213> Homo sapiens
<400> 98
Met Ala Ser Gly Asp Leu Phe Pro Gly Asn Val Phe Ser Pro Arg Val
                                    10
Pro Phe Pro Ser Ser Asp Ile Thr Pro Asn Ile Glu Arg Leu Ser Asn
Ser Pro Leu Arg Ser Ser Arg Ala Tyr Ala Gln Asp Phe Met Thr Leu
                            40
Asp Lys Arg Trp His Val Met Ala Gly Ile Ala Leu Asn Gln Leu Pro
                        55
    50
Gln Glu Gly Gly Pro Thr Glu Arg Ala Trp Thr Pro Lys Leu Gly Leu
                                        75
                    70
Ser Phe Asp Val Ser Asp Thr Met Ser Leu Tyr Gly Ala Tyr Ser Arg
                                    90
Gly Phe Ser Thr Tyr Gln Pro Ala Arg Lys Ala Pro Arg Ala Tyr Gly
            100
                                105
                                                     110
Pro Ser Ala Ala Arg Pro Ser Lys Arg Glu
                            120
        115
<210> 99
<211> 545
<212> DNA
<213> Homo sapiens
<400> 99
acgegteege tecegatgte gttgacgage tgegeteage gatgacggta eteacecate
tgccccgacg acccagcaaa cgtccccggc tgttcctcat tgaccacgcc gaccggatcg
togatoccat cactogggat ttgctggaat coctggttcg cgaagccggc gaggctgcgg
tgatcttggg tgcccagcgt cgcggtcgca tcgattggct ctccccacag atcatccaca
240
```

```
acctggccga acaccatttt gagtcgtcct ctggaggtac tcgatgatga ctgaacgttc
ccattccacg atcaggttaa ggtggccggc ggtggtggtt ctcgtcctcg ttccgctgct
ggtggtcgcc ggattggtcc gggacgacct ggcataccac cgaccggttg ggccgggtga
420
aageggeegt egteaaegag gacaaggeeg teaaggtgeg tggacaaetg gtteegatgg
480
geegecaact caeegeegee ttgatggaet etggetegea caeeactgat ggeeacaeeg
540
tcgac
545
<210> 100
<211> 101
<212> PRT
<213> Homo sapiens
<400> 100
Met Gly Thr Phe Ser His His Arg Val Pro Pro Glu Asp Asp Ser Lys
                                                         15
                                    10
Trp Cys Ser Ala Arg Leu Trp Met Ile Cys Gly Glu Ser Gln Ser Met
                                 25
Arg Pro Arg Arg Trp Ala Pro Lys Ile Thr Ala Ala Ser Pro Ala Ser
                             40
Arg Thr Arg Asp Ser Ser Lys Ser Arg Val Met Gly Ser Thr Ile Arg
                        5.5
Ser Ala Trp Ser Met Arg Asn Ser Arg Gly Arg Leu Leu Gly Arg Arg
                                         75
Gly Arg Trp Val Ser Thr Val Ile Ala Glu Arg Ser Ser Ser Thr Thr
                                     90
                 85
Ser Gly Ala Asp Ala
            100
<210> 101
 <211> 619
 <212> DNA
 <213> Homo sapiens
 <400> 101
ngegegecae agaagtgaaa aagtaeggte teaagecegt gggagegtte tteetgteeg
 60
 teatgeacga teeggaette gateegatee eeatggtgaa caaggagett gaegeetteg
 aagetgeegg gggtgactat eteateeteg eeaeggatte eggaegeaag ggatacaega
 cegecegtee teacgaggee ggeggaaaac getattacca acetggacca gateegegaa
 gtetgegeca geogeaacgt cacegootgt ctacaccooc attggggaac gatggtecag
 aaccgtgacg aagtgatccg cgtgctcgag aactcctcga tcgggctgtg cctggacact
 ggtcatctgg cctgtggtgg taccgatgtc gttgagctgg tgcgtaagta cgccaaccgc
 420
```

```
gtcgacattg tccacgccaa agatgtccat aaggagatgg ccgacaagct tttgcctggc
480
gagateacet ggteegaagg cattegegee gggatgtteg eacecategg egaeggtgat
ategaetttg cagecategt gaggeteett gatgaageeg ggttegatgg ttattaegte
ctagagcagg acatcatga
619
<210> 102
<211> 173
<212> PRT
<213> Homo sapiens
<400> 102
Thr Arg Ser Leu Thr Pro Ser Lys Leu Pro Gly Val Thr Ile Ser Ser
                                   10
Ser Pro Arg Ile Pro Asp Ala Arg Asp Thr Arg Pro Pro Val Leu Thr
                                25
Arg Pro Ala Glu Asn Ala Ile Thr Asn Leu Asp Gln Ile Arg Glu Val
                            40
Cys Ala Ser Arg Asn Val Thr Ala Cys Leu His Pro His Trp Gly Thr
                                            60
                        55
Met Val Gln Asn Arg Asp Glu Val Ile Arg Val Leu Glu Asn Ser Ser
                    70
Ile Gly Leu Cys Leu Asp Thr Gly His Leu Ala Cys Gly Gly Thr Asp
                                    90
Val Val Glu Leu Val Arg Lys Tyr Ala Asn Arg Val Asp Ile Val His
                                105
            100
Ala Lys Asp Val His Lys Glu Met Ala Asp Lys Leu Leu Pro Gly Glu
                            120
        115
Ile Thr Trp Ser Glu Gly Ile Arg Ala Gly Met Phe Ala Pro Ile Gly
                        135
Asp Gly Asp Ile Asp Phe Ala Ala Ile Val Arg Leu Leu Asp Glu Ala
                                        155
                    150
Gly Phe Asp Gly Tyr Tyr Val Leu Glu Gln Asp Ile Met
                                    170
                165
<210> 103
<211> 321
<212> DNA
<213> Homo sapiens
<400> 103
nnccatgggg gaagacaaca gccatgtggg ggagacccga gccattgggg ggagacccct
gecattgggg ggagacecet geegtgggga aagaceeetg ceatggggea gaeeeetgee
actgggggga gacccctgcc gctgggggga gacccgagcc attgggggga gacccctgcc
atggggaaag acccctgcca ttgggggaga ntacctgcca ttgggggaga tccctgccgt
tggggggaga tccctgctgt tggggggaga ntccctcctg taggggaaga cccctgcagg
300
```

```
agtggttggg gcgaagaccc c
<210> 104
<211> 107
<212> PRT
<213> Homo sapiens
<400> 104
Xaa His Gly Gly Arg Gln Gln Pro Cys Gly Gly Asp Pro Ser His Trp
                                    10
1
Gly Glu Thr Pro Ala Ile Gly Gly Arg Pro Leu Pro Trp Gly Lys Thr
                                                     3.0
                                25
Pro Ala Met Gly Gln Thr Pro Ala Thr Gly Gly Arg Pro Leu Pro Leu
                            40
Gly Gly Asp Pro Ser His Trp Gly Glu Thr Pro Ala Met Gly Lys Asp
Pro Cys His Trp Gly Arg Xaa Pro Ala Ile Gly Gly Asp Pro Cys Arg
                                        75
                    70
Trp Gly Glu Ile Pro Ala Val Gly Gly Arg Xaa Pro Pro Val Gly Glu
Asp Pro Cys Arg Ser Gly Trp Gly Glu Asp Pro
            100
<210> 105
<211> 344
<212> DNA
<213> Homo sapiens
<400> 105
nnntetetec gacegegtee agatecaceg tggcccgcat gaaccagteg ttgttgeett
ccgggtcaac gagggtttgg cgcacggtcc actccgtggc gcccggggtg atgtgcaaca
120
gggcgggccc gcgcgcggcc gggcctgatt ccagcctctc gtgctcgtcc cagtacccat
ccagegeate gecceagegg teggeatece ageegtegte geegtegage geceeeaggg
240
cotcaatgto gtoatoggog gooagttoca cooggoggaa catotogttg oggacoatga
cccggaaggc gcgggaattc tcggtcagtt tcggcggtgc cggc
344
<210> 106
<211> 62
<212> PRT
<213> Homo sapiens
<400> 106
Cys Ala Thr Gly Arg Ala Arg Ala Arg Pro Gly Leu Ile Pro Ala Ser
                                     10
Arg Ala Arg Pro Ser Thr His Pro Ala His Arg Pro Ser Gly Arg His
                                 25
Pro Ser Arg Gly Arg Arg Arg Ala Pro Pro Gly Pro Gln Cys Arg His
```

```
40
Arg Arg Pro Val Pro Pro Gly Gly Thr Ser Arg Cys Gly Pro
<210> 107
<211> 549
<212> DNA
<213> Homo sapiens
<400> 107
acgcgttcct cggacgctgc tagtttgagc gcacaggcat tgggcggggc tcacggggga
geogettaat aacegaceaa catgaaacte aagggetgee cetteetage ggggaceetg
cacagacccg aaaataaggg gttttgctct gccctcctca gttcacgtgg gcaccttgga
acactgaaga aggeatttte egaacteaet gteetaegga ettattetee geactgtttt
cgcctccttc gccctgttct cgtgactgac aggagcaggg gtcacaagca ggcagcccga
qaqctctqct cacctggaaa agcatttttg tgtagcttaa atgtgaaggc ctcaggcagt
ggcctgttgt cctcctccac atgcgcccat cttcactctt tcatgtgact ggcctgtttt
tgaaggcaag gcccctgtca cccttggcta ggccaggtat gttctgcacc gaaaatggcc
ctgccctctg cattggatgg ctagctctta ggttggttta ttttagcaaa taagcgttac
540
agggtaggc
549
<210> 108
<211> 108
<212> PRT
<213> Homo sapiens
<400> 108
Met Lys Leu Lys Gly Cys Pro Phe Leu Ala Gly Thr Leu His Arg Pro
                                    10
Glu Asn Lys Gly Phe Cys Ser Ala Leu Leu Ser Ser Arg Gly His Leu
                                25
Gly Thr Leu Lys Lys Ala Phe Ser Glu Leu Thr Val Leu Arg Thr Tyr
Ser Pro His Cys Phe Arg Leu Leu Arg Pro Val Leu Val Thr Asp Arg
Ser Arg Gly His Lys Gln Ala Ala Arg Glu Leu Cys Ser Pro Gly Lys
Ala Phe Leu Cys Ser Leu Asn Val Lys Ala Ser Gly Ser Gly Leu Leu
                85
Ser Ser Ser Thr Cys Ala His Leu His Ser Phe Met
            100
                                105
<210> 109
<211> 748
```

<212> DNA <213> Homo sapiens <400> 109 nngaattcag atttactttt tgcatttcct tgaatataaa ataggcagta aactaaccta aagtcagaag acttatttgt totggtacot attotgotgt taataaattg ataaatgagt gatattgggg taagttacca atttacttta cagcccttaa gtaaataatc tgctttcctc aqcatcatag actittgaag aggattaatt aagcgcttaa aaaacctgta gactctatta 240 cagtcagtga aaggaataat tototttaca aagtaaatgo agttgtttta ttttagacaa 300 gagtgttcta aacttcgtga agagttaagg cttcaacatg aagaggataa gaagtcagca atgtctcaac ttttgcagtt gaaagatcga gagaaaaatg cagcaagaga ttcatggcag aagaaagtag aagatetett aaaccagatt teettgetga aacagaatet ggagatacag ctttcccagt ctcagacttc tttgcaacaa ctgcaagccc agtttacgca agaacgacag cggcttacgc aagagcttga agaattagag gagcaacatc agcaaagaca caaatcatta aaagaagcac atgtccttgc atttcaaact atggaagagg aaaaggaaaa ggagcaaaga gctcttgaaa atcatttaca acagaagcat tctgcagagc ttcaatcact aaaagatgca cacagagagt caatggaggg cttccgga 748 <210> 110 <211> 157 <212> PRT <213> Homo sapiens <400> 110 Met Gln Leu Phe Tyr Phe Arg Gln Glu Cys Ser Lys Leu Arg Glu Glu Leu Arg Leu Gln His Glu Glu Asp Lys Ser Ala Met Ser Gln Leu Leu Gln Leu Lys Asp Arg Glu Lys Asn Ala Ala Arg Asp Ser Trp Gln Lys Lys Val Glu Asp Leu Leu Asn Gln Ile Ser Leu Leu Lys Gln Asn 55 Leu Glu Ile Gln Leu Ser Gln Ser Gln Thr Ser Leu Gln Gln Leu Gln 70 75 Ala Gln Phe Thr Gln Glu Arg Gln Arg Leu Thr Gln Glu Leu Glu Glu 90 Leu Glu Gln His Gln Gln Arg His Lys Ser Leu Lys Glu Ala His 105 Val Leu Ala Phe Gln Thr Met Glu Glu Lys Glu Lys Glu Gln Arg 120 Ala Leu Glu Asn His Leu Gln Gln Lys His Ser Ala Glu Leu Gln Ser

```
135
    130
Leu Lys Asp Ala His Arg Glu Ser Met Glu Gly Phe Arg
145
                    150
<210> 111
<211> 429
<212> DNA
<213> Homo sapiens
<400> 111
gcgcgcccag agagtgcaca atggtgtcag gatatgggag ccaccggaat cattaaccat
cgtgaaccgc tcgccccgca ggtcaacgac ttcgggatca ccgggttcga cggcattctc
teggettatn nacgecaeca geatnegaet ttggetgaga teategeaec gtteggaeat
180
ctggtcatga tcgacggaac cgactcattc gatctcatgg ccttcaagtc aaagtcgtta
acggtgacca gcgagtcgat gttcagccgt ccacagttcg cgacgcccga cgtcgccgaa
caaggeeggg caetggeeag categeegae etegtegaga aggggeagat eegteegaeg
atgaccogcc atatogaggg totgacaacc cagcatgtgc gtgaggccac cgcagccgtc
420
gagtccggc
429
<210> 112
<211> 143
<212> PRT
<213> Homo sapiens
<400> 112
Ala Arg Pro Glu Ser Ala Gln Trp Cys Gln Asp Met Gly Ala Thr Gly
                                    10
1
Ile Ile Asn His Arg Glu Pro Leu Ala Pro Gln Val Asn Asp Phe Gly
            20
                                25
Ile Thr Gly Phe Asp Gly Ile Leu Ser Ala Tyr Xaa Arg His Gln His
                            40
Xaa Thr Leu Ala Glu Ile Ile Ala Pro Phe Gly His Leu Val Met Ile
Asp Gly Thr Asp Ser Phe Asp Leu Met Ala Phe Lys Ser Lys Ser Leu
                                        75
Thr Val Thr Ser Glu Ser Met Phe Ser Arg Pro Gln Phe Ala Thr Pro
                85
Asp Val Ala Glu Gln Gly Arg Ala Leu Ala Ser Ile Ala Asp Leu Val
                                105
Glu Lys Gly Gln Ile Arg Pro Thr Met Thr Arg His Ile Glu Gly Leu
                            120
                                                125
Thr Thr Gln His Val Arg Glu Ala Thr Ala Ala Val Glu Ser Gly
    130
                        135
<210> 113
<211> 382
```

<212> DNA

```
<213> Homo sapiens
<400> 113
gtgcactgga cggatgctgg ggaactgagg cctgcccttc tccactcctg ctgcctctgt
gtggaccgcc tgccacctgg cttcaacgat gtggacgctc tgtgccgggc gctgtcagct
120
gtccacagcc ccaccttctg ccagctggcg tgcggccagg atgggcagct gaagggcttc
180
geggtgetgg agtatgagac ggetgagatg geggaggagg cacageagea ggeggaegge
240
ctgtccctgg ggggcagcca cctgcgagtc tccttctgcg cccctgggac ccccgggcgc
agtatgetgg cegeteteat egttgeecag gecaeggeee teaategggg geagggagte
ctccccgagc ccaacatcct gc
382
<210> 114
<211> 125
<212> PRT
<213> Homo sapiens
Met Leu Gly Ser Gly Arg Thr Pro Cys Pro Arg Leu Arg Ala Val Ala
                                    10
Trp Ala Thr Met Arg Ala Ala Ser Ile Leu Arg Pro Gly Val Pro Gly
                                25
            20
Ala Gln Lys Glu Thr Arg Arg Trp Leu Pro Pro Arg Asp Arg Pro Ser
                            40
Ala Cys Cys Cys Ala Ser Ser Ala Ile Ser Ala Val Ser Tyr Ser Ser
                                            60
                        55
Thr Ala Lys Pro Phe Ser Cys Pro Ser Trp Pro His Ala Ser Trp Gln
                    70
                                        75
Lys Val Gly Leu Trp Thr Ala Asp Ser Ala Arg His Arg Ala Ser Thr
                                    90
Ser Leu Lys Pro Gly Gly Arg Arg Ser Thr Gln Arg Gln Gln Glu Trp
                                105
Arg Arg Ala Gly Leu Ser Ser Pro Ala Ser Val Gln Cys
                                                 125
                            120
<210> 115
<211> 4798
<212> DNA
<213> Homo sapiens
<400> 115
ctcccttcca gtctgtctta agtactgatg ccagagaggt ctgcctaagc tcagtgctgt
catttcactc cctgtttgga gccatgagtg gccccctgtt gccttcaaga ctgaagccat
coccatoctt cotcocacca acctagagge tttgettegt aaatgetgge cettteette
```

atgggctcca 240	ccctctgagt	gtgtcatgaa	ctcttccact	teettgeett	ggttcgtatt
gggtcctctg 300	cccgaggtca	gagatttgga	cgagcccttc	tcctccatct	tcacagtctc
catttaccat 360	tagatgccag	gcccagacag	ttgatgaaaa	atgtaagaaa	cggacccaag
tataagaaaa 420	agctagatct	gtccatttta	tttctagttt	ttggaagaga	tgtggataca
gagaaatatt 480	actctactat	tgaaaaaaat	atatatcagt	ctgattataa	atgtcaactc
atactaagtc 540	aatatcaggg	aagcagtgtg	gagaaggatg	tggcaaactg	gagattgctg
tttccataga 600	aagagggcag	ctaatgctta	tctccagacc	attgttatgc	agtaatgtag
actccatttc 660	accagattta	attcttaagg	agaaactgga	aacttgtatt	attatttgat
tattactttt 720	ggcagttaat	tttaaaaatc	accgtttagg	ccaaaacaac	atgacatttg
tctgggctac 780	aagggactct	tagcgactct	tagcttgcac	ttacctttct	aatctcaggg
ctttttacaa 840	agcccttctc	tgaattccct	gaagatatgc	tagcttgaag	gctcactgca
tactctcaac 900	ttcccagcca	cttccctgtt	tccctcttat	accagttaat	actttgatgt
tattttttca 960	gtcctattta	atcttttaga	ctggaggttc	cttgagggtg	gggcctaatt
cttgttttta 1020	tettececat	agcaattatc	aagttatctg	ctccctttgg	cacctaagaa
atattagttg 1080	gggtgataac	tgaatggctg	tecttetgee	ttcgttactg	cttttgtgga
tgttattttc 1140	tttgacccct	attccacccg	tgtgcagtgg	agaggatggg	tgttgtattt
tacagaggaa 1200	ggaattaact	gagacgcaac	aggatgatgg	gaacttctca	aggtcatgtg
gcaaggaagt 1260	caaggaattg	gggtttgaac	ccaagtcgtc	tcagttccat	tccattgtcc
tctaccccct	gccacctctc	accttcttcc	ctttcacctt	tttctgtagc	tgaacggaaa
ccaccgctct 1380	ttaacatgaa	tgcgatgagt	gccttatacc	acattgcaca	gaacgaatcc
cccgtgctcc	agtcaggaca	ctggtctgag	tacttccgga	attttgtcga	ctcctgtctt
cagaaaatcc 1500	ctcaagacag	accaacctca	gaggttctcc	tgaagcaccg	ctttgtgctc
cgggagcggc 1560	cacccacagt	catcatggac	ctgatccaga	ggaccaagga	tgccgtgcgg
gagctggaca 1620	acctgcagta	ccgcaagatg	aagaagatcc	tgttccaaga	ggcacccaac
ggccctggtg 1680	ccgaggcccc	agaggaggaa	gaggaggccg	agccctacat	gcaccgggcc
gggactctga 1740	ccagcctcga	gagtagccac	tcagtgccca	gcatgtccat	cagcgcctcc
agccagagca 1800	gctccgtcaa	cagectagea	gatgcctcag	acaacgagga	agaggaggag

	aagaggagga	ggaggaagaa	ggccctgaag	cccgggagat	ggccatgatg
1860 caggagggg	agcacacagt	cacctctcac	agctccatta	tccaccggct	gccgggctct
1920	atgatgaccc	ctaccagcca	gagataaccc	ccagccctct	ccagccgcct
1980					
gcagccccag 2040	ctcccacttc	caccacctct	teegeeegee	gccgggccta	ctgccgtaac
cgagaccact 2100	ttgccaccat	ccgaaccgcc	tccctggtca	gccgtcagat	ccaggagcat
gagcaggact 2160	ctgcgctgcg	ggagcagctg	agcggctata	ageggatgeg	acgacagcac
	tgctggccct	ggagtcacgg	ctgaggggtg	aacgggagga	gcacagtgca
	gggagcttga	ggcgcagcgg	gctggctttg	gggcagaggc	agaaaagctg
	accaggccat	aggtgagaag	gaggcacgag	ctgcccaggc	cgaggagcgg
	agcacatcct	tgggcagcag	aagaaggagc	tggctgccct	gctggaggca
	cctacaaact	tcgcaaggaa	cagetgaagg	aggagctcca	ggagaacccc
	agcgggagaa	ggccgagtgg	ctgctgcggc	agaaggagca	gctccagcag
	aggaggaagc	agggetgetg	cggcggcagc	gccagtactt	tgagctgcag
	acaagcgcaa	gatgttgctg	gctcggcaca	gcctggacca	ggacctgctg
	tgaacaagaa	gcagacccag	aaggacttgg	agtgtgcact	gctgcttcgg
cagcacgagg 2760	ccacgcggga	gctggagctg	cggcagctcc	aggccgtgca	gcgcacgcgg
gctgagctca 2820	cccgcctgca	gcaccagacg	gagetgggca	accagetgga	gtacaacaag
cggcgtgagc 2880	aagagttgcg	gcagaagcat	geggeeeagg	ttcgccagca	geceaagage
ctcaaagtac 2940	gtgcaggcca	gegeeeeeeg	ggccttccac	tccccattcc	tggggctctg
	acacaggcac	ccctatagaa	cagcagccct	gctcacctgg	ccaggaggca
gtcctggacc 3060	aaagaatgct	tggcgaggag	gaggaagcag	ttggagagag	aaggattctg
ggaaaggaag 3120	gggccacttt	ggagcccaag	cagcagagga	ttctggggga	agaatcagga
gcccctagtc 3180	ccagtccaca	aaaacatggg	agcctggttg	atgaggaagt	ttggggtctg
	tagaggagct	tagggtgccc	tcccttgtac	cccaggagag	gagcattgtt
	aggctgggac	gtggagcttg	tgggggaagg	aggatgagag	tcttctggat
	agcttggctg	ggtccagggc	ccagcactga	ctcccgtccc	tgaggaggag
	aagaggggc	tccgattggg	acccctaggg	atcctggaga	tggttgtcct

```
tecceegaca tecetectga acceeeteca acacaeetga ggeeetgeee tgeeageeag
ctccctqqac tcctqtccca tggcctcctg gccggcctct cctttgcagt ggggtcctcc
tetggeetee tgeeceteet getgetgetg etgetteeat tgetggeage ceagggtggg
ggtggcctgc aggcagcgct gctggccctt gaggtggggc tggtgggtct gggggcctcc
tacctgctcc tttgtacagc cctgcacctg ccctccagtc ttttcctact cctggcccag
3720
ggtaccgcac tgggggccgt cctgggcctg agctggcgcc gaggcctcat gggtgttccc
3780
ctgggccttg gagctgcctg gctcttagct tggccaggcc tagctctacc tctggtggct
atggcagcgg ggggcagatg ggtgcggcag caaggccccc gggtgcgccg gggcatatct
cgactetggt tgcgggttet gctgcgcctg tcacccatgg ccttccgggc cctgcagggc
3960
tqtqgggctg tgggggaccg ggggctgttt gcactgtacc ccaaaaccaa caaggatggc
4020
ttecqcagec geetgeeegt ceetgggeee eggeggegta ateceegeae caeceaacae
ccattagete tgttggcaag ggtctgggte ctgtgcaagg gctggaactg gcgtctggca
cgggccagcc agggtttagc atcccacttg cccccgtggg ccatccacac actggccagc
tggggcctgc ttcggggtga acggcccacc cgaatccccc ggctactacc acgcagccag
cgccagctgg ggccccctgc ctcccgccag ccactgccag ggactctagc cgggcggagg
teacgeacce gecagteceg ggecetgece ecetggaggt agetgactee agecetteca
geccaaatet agageattga geaetttate teecacgaet cagtgaagtt tetecagtee
ctagtcctct cttttcaccc accttcctca gtttgctcac ttaccccagg cccagccctt
cggacctcta gacaggcage etecteaget gtggagtcca gcagtcactc tgtgttetec
tggcgctcct cccctaagtt attgctgttc gcccgctgtg tgtgctcatc ctcaccctca
4620
ttgactcagg cctggggcca ggggtggtgg agggtgggaa gagtcatgtt ttttttctcc
tetttgattt tgtttttetg tetceettee aacetgteee etteeecea ecaaaaaaag
4798
<210> 116
<211> 1062
<212> PRT
<213> Homo sapiens
<400> 116
Met Met Gly Thr Ser Gln Gly His Val Ala Arg Lys Ser Arg Asn Trp
```

,				_					10					15	
l Gly	Leu	Δen	Dro	5 Ser	Ara	T. 211	Ser	Ser		Pro	T.011	Ser	Ser	-	Pro
Gry	пец	ASII	20	261	Arg	пеа	SCI	25	116	FIO	пец	Jer	30	1111	110
Cvs	His	Leu		Pro	Ser	Ser	Leu		Pro	Phe	Ser	Val		Glu	Ara
Cyb		35	501		001		40	-				45			
Lvs	Pro		Leu	Phe	Asn	Met		Ala	Met	Ser	Ala		Tyr	His	Ile
-7-	50					55					60		- 4 -		
Ala	Gln	Asn	Glu	Ser	Pro	Val	Leu	Gln	Ser	Gly	His	Trp	Ser	Glu	Tyr
65					70					75		_			80
Phe	Arg	Asn	Phe	Val	Asp	Ser	Cys	Leu	Gln	Lys	Ile	Pro	Gln	Asp	Arg
	_			85					90					95	
Pro	Thr	Ser	Glu	Val	Leu	Leu	Lys	His	Arg	Phe	Val	Leu	Arg	Glu	Arg
			100					105					110		
Pro	Pro	Thr	Val	Ile	Met	Asp	Leu	Ile	Gln	Arg	Thr	Lys	Asp	Ala	Val
	_	115					120					125		_	
Arg	Glu	Leu	Asp	Asn	Leu		Tyr	Arg	Lys	Met		Lys	Ile	Leu	Phe
~1	130		_		~1	135	~ 1		~1		140	~1	~1	a 1	a 1
	Glu	Ala	Pro	Asn		Pro	GIY	Ala	GIU		Pro	GIU	GIU	GIU	160
145	Ala	~1.,	Dwo	Т. г.	150	TI i a	7. ~~~	- ו מ	C1	155	Lou	The	co-	Lou	
GIU	ALA	Giu	PIO	165	MEC	піз	Arg	нта	170	1111	Leu	1111	261	175	GIU
Ser	Ser	His	Ser		Pro	Ser	Met	Ser		Ser	Δla	Ser	Ser		Ser
DCI	DCI		180	var	110	001		185	110				190		501
Ser	Ser	Val		Ser	Leu	Ala	Asp		Ser	Asp	Asn	Glu		Glu	Glu
		195					200			•		205			
Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Gly	Pro	Glu	Ala	Arg
	210					215					220				
Glu	Met	Ala	Met	Met	Gln	Glu	Gly	Glu	His	Thr	Val	Thr	Ser	His	Ser
225					230					235					240
Ser	Ile	Ile	His		Leu	Pro	Gly	Ser		Asn	Leu	Tyr	Asp		Pro
				245		_	_	_	250		_	_		255	_
Tyr	Gln	Pro		Ile	Thr	Pro	Ser		Leu	GIn	Pro	Pro		Ala	Pro
21-	D	The see	260	Th se	mb se	C	C	265	7	7	7	7 T -	270	Cira	7 ~~~
Ата	Pro	275	ser	Int	THE	Ser	280	Ala	Arg	Arg	Arg	285	TÄT	Cys	Arg
Λen	Arg		Hic	Phe	Δla	Thr		Δrα	Thr	Δla	Ser		val	Ser	Ara
Maii	290	ASP	1113	1110	n.a	295	110	nr 9	1111	Aru	300		var		9
Gln	Ile	Gln	Glu	His	Glu		Asp	Ser	Ala	Leu		Glu	Gln	Leu	Ser
305		•	•		310		•			315					320
	Tyr	Lys	Arg	Met	Arg	Arg	Gln	His	Gln	Lys	Gln	Leu	Leu	Ala	Leu
-	-	_	_	325	_	_			330	_				335	
Glu	Ser	Arg	Leu	Arg	Gly	Glu	Arg	Glu	Glu	His	Ser	Ala	Arg	Leu	Gln
			340					345					350		
Arg	Glu	Leu	Glu	Ala	Gln	Arg	Ala	Gly	Phe	Gly	Ala	Glu	Ala	Glu	Lys
		355					360					365		_	_
Leu	Ala	Arg	Arg	His	Gln		Ile	Gly	Glu	Lys		Ala	Arg	Ala	Ala
	370				_	375			•		380			~1	_
	Ala	GIu	GIU	Arg	_	Pne	GIN	GIN	HIS		Leu	GTĀ	GIN	GIN	_
385	~1	T =	 - ד ת	71-	390	T 6	<i>c</i> 1	7 7 -	01 -	395	7	rmb so	TT a see	T~	400
ьуѕ	Glu	ьeu	нта	405	reu	ьeu	GIU	нта		гÀг	Arg	TILL	TAT		ьeu
7/~~	Lys	Glas	Gln		Live	Glu	Glu	Len	410 Gln	Glu	Aen	Pro	Ser	415 Thr	Dro
719	пyэ	Jiu	420	u u	-13	JIU	Jau	425	-11	Jau	II		430	****	210
Lve	Arg	Glu		Ala	Glu	Trp	Leu		Ara	GÌn	Lvs	Glu		Leu	Gln
_, _	5		-1-			E			5		-1-				

												445			
Gln (~	435	.1.	~1	~1	~1	440	C1	T 011	T ON	7 20		Gl n	Δνα	Gln
	.ys 450	GIN	АІА	GIU	GIU	455	ALA	GIY	Leu	Leu	460	Arg	GIII	Arg	0111
Tyr E		Glu	T 011	Gln	Cve		Gln	Tur	Lve	Δrσ		Met	Leu	Leu	Ala
465	PHE	GIU	ьец	GIII	470	nr 9	GIII	- y -	L y S	475	_,_			200	480
Arg F	Hie	Ser	Leu	Asn		Asp	Leu	Leu	Ara	-	Asp	Leu	Asn	Lvs	
Arg I	113	561	LCu	485	U111	1100			490					495	_1
Gln T	Thr	Gln	Lvs		Leu	Glu	Cvs	Ala		Leu	Leu	Ara	Gln	His	Glu
GIII .		Q	500	p	200		010	505					510		
Ala T	Thr	Ara		Leu	Glu	Leu	Ara		Leu	Gln	Ala	Val		Arq	Thr
7114	- 	515					520					525		_	
Arg A	Ala		Leu	Thr	Arq	Leu	Gln	His	Gln	Thr	Glu	Leu	Gly	Asn	Gln
	530					535					540		-		
Leu G		Tyr	Asn	Lys	Arg	Arg	Glu	Gln	Glu	Leu	Arg	Gln	Lys	His	Ala
545		•		-	550	_				555					560
Ala G	Gln	Val	Arg	Gln	Gln	Pro	Lys	Ser	Leu	Lys	Val	Arg	Ala	Gly	Gln
				5 65					570					575	
Arg I	Pro	Pro	Gly	Leu	Pro	Leu	Pro	Ile	Pro	Gly	Ala	Leu	Gly	Pro	Pro
			580					5 85					590		
Asn 7	Thr	Gly	Thr	Pro	Ile	Glu	Gln	Gln	Pro	Cys	Ser		Gly	Gln	Glu
		595					600					605	_		
Ala \	Val	Leu	Asp	Gln	Arg	Met	Leu	Gly	Glu	Glu		Glu	Ala	Val	Gly
	610				_	615	_				620		_	_	~3
Glu A	Arg	Arg	Ile	Leu		Lys	Glu	Gly	Ala		Leu	GLu	Pro	гĀг	
625			_		630	~-3	_	~1		635	a	D		Dage	640
Gln A	Arg	Ile	Leu		GLu	Glu	Ser	GLY		Pro	ser	Pro	ser	655	GIII
Lys F		~1	C	645	1707	7.00	C1	C1	650	Trn	Glv	T 211	Dro		Glu
rys i	HIS	GIY	660	Leu	val	Asp	GIU	665	Val	TTD	GIY	шец	670	Giu	GIU
Ile (C1	Clu		7 200	17 - 1	Dro	Sar		พลใ	Pro	Gln	Glu		Ser	Tle
TIE (Giu	675	пеп	AL 9	vai	110	680	LCu	Val		U	685	5		
Val (Glv		Glu	Glu	Δla	Glv		Trp	Ser	Leu	Trp		Lvs	Glu	Asp
	690	0111	Olu	0.10		695					700	4	- 4		-
Glu s		Leu	Leu	Asp	Glu		Dhe	01	_	Cl.	Tro	Val	Gln	Glv	Pro
705							LIIC	$u_{\perp u}$	Leu	GTA					
Ala I	_			_	710	014	rne	GIU	Leu	715				0-1	720
	Leu	Thr	Pro		710					715					
	Leu	Thr	Pro		710					715					
Pro 1				Val 725	710 Pro	Glu	Glu	Glu	Glu 730	715 Glu	Glu	Glu	Glu	Gly 735	Ala
	Ile	Gly	Thr 740	Val 725 Pro	710 Pro Arg	Glu Asp	Glu Pro	Glu Gly 745	Glu 730 Asp	715 Glu Gly	Glu Cys	Glu Pro	Glu Ser 750	Gly 735 Pro	Ala Asp
Pro I	Ile	Gly	Thr 740	Val 725 Pro	710 Pro Arg	Glu Asp	Glu Pro	Glu Gly 745	Glu 730 Asp	715 Glu Gly	Glu Cys	Glu Pro	Glu Ser 750	Gly 735 Pro	Ala Asp
Ile I	Ile Pro	Gly Pro 755	Thr 740 Glu	Val 725 Pro	710 Pro Arg Pro	Glu Asp Pro	Glu Pro Thr 760	Glu Gly 745 His	Glu 730 Asp Leu	715 Glu Gly Arg	Glu Cys Pro	Glu Pro Cys 765	Glu Ser 750 Pro	Gly 735 Pro Ala	Ala Asp Ser
	Ile Pro	Gly Pro 755	Thr 740 Glu	Val 725 Pro	710 Pro Arg Pro	Glu Asp Pro	Glu Pro Thr 760	Glu Gly 745 His	Glu 730 Asp Leu	715 Glu Gly Arg	Glu Cys Pro Ala	Glu Pro Cys 765	Glu Ser 750 Pro	Gly 735 Pro Ala	Ala Asp Ser
Ile I	Ile Pro Leu 770	Gly Pro 755 Pro	Thr 740 Glu Gly	Val 725 Pro Pro	710 Pro Arg Pro Leu	Glu Asp Pro Ser 775	Glu Pro Thr 760 His	Glu Gly 745 His	Glu 730 Asp Leu Leu	715 Glu Gly Arg Leu	Glu Cys Pro Ala 780	Glu Pro Cys 765 Gly	Glu Ser 750 Pro Leu	Gly 735 Pro Ala Ser	Ala Asp Ser Phe
Ile I	Ile Pro Leu 770	Gly Pro 755 Pro	Thr 740 Glu Gly	Val 725 Pro Pro	710 Pro Arg Pro Leu Ser	Glu Asp Pro Ser 775	Glu Pro Thr 760 His	Glu Gly 745 His	Glu 730 Asp Leu Leu	715 Glu Gly Arg Leu Leu	Glu Cys Pro Ala 780	Glu Pro Cys 765 Gly	Glu Ser 750 Pro Leu	Gly 735 Pro Ala Ser	Ala Asp Ser Phe Leu
Ile II Gln II Ala V	Ile Pro Leu 770 Val	Gly Pro 755 Pro Gly	Thr 740 Glu Gly Ser	Val 725 Pro Pro Leu Ser	710 Pro Arg Pro Leu Ser 790	Glu Asp Pro Ser 775 Gly	Glu Pro Thr 760 His	Glu Gly 745 His Gly Leu	Glu 730 Asp Leu Leu Pro	715 Glu Gly Arg Leu Leu 795	Glu Cys Pro Ala 780 Leu	Glu Pro Cys 765 Gly Leu	Glu Ser 750 Pro Leu Leu	Gly 735 Pro Ala Ser Leu	Ala Asp Ser Phe Leu 800
Ile H	Ile Pro Leu 770 Val	Gly Pro 755 Pro Gly	Thr 740 Glu Gly Ser	Val 725 Pro Pro Leu Ser	710 Pro Arg Pro Leu Ser 790	Glu Asp Pro Ser 775 Gly	Glu Pro Thr 760 His	Glu Gly 745 His Gly Leu	Glu 730 Asp Leu Leu Pro	715 Glu Gly Arg Leu Leu 795	Glu Cys Pro Ala 780 Leu	Glu Pro Cys 765 Gly Leu	Glu Ser 750 Pro Leu Leu	Gly 735 Pro Ala Ser Leu Ala	Ala Asp Ser Phe Leu 800
Ile II Gln II Ala V 785 Leu I	Ile Pro Leu 770 Val	Gly Pro 755 Pro Gly Leu	Thr 740 Glu Gly Ser Leu	Val 725 Pro Pro Leu Ser Ala 805	710 Pro Arg Pro Leu Ser 790 Ala	Glu Asp Pro Ser 775 Gly Gln	Glu Pro Thr 760 His Leu Gly	Glu Gly 745 His Gly Leu Gly	Glu 730 Asp Leu Leu Pro Gly 810	715 Glu Gly Arg Leu Leu 795 Gly	Glu Cys Pro Ala 780 Leu	Glu Pro Cys 765 Gly Leu Gln	Glu Ser 750 Pro Leu Leu	Gly 735 Pro Ala Ser Leu Ala 815	Ala Asp Ser Phe Leu 800 Leu
Ile II Gln II Ala V	Ile Pro Leu 770 Val	Gly Pro 755 Pro Gly Leu	Thr 740 Glu Gly Ser Leu Glu	Val 725 Pro Pro Leu Ser Ala 805	710 Pro Arg Pro Leu Ser 790 Ala	Glu Asp Pro Ser 775 Gly Gln	Glu Pro Thr 760 His Leu Gly	Glu Gly 745 His Gly Leu Gly Gly	Glu 730 Asp Leu Leu Pro Gly 810	715 Glu Gly Arg Leu Leu 795 Gly	Glu Cys Pro Ala 780 Leu	Glu Pro Cys 765 Gly Leu Gln	Glu Ser 750 Pro Leu Leu Ala Tyr	Gly 735 Pro Ala Ser Leu Ala 815	Ala Asp Ser Phe Leu 800 Leu
Ile II Gln II Ala V 785 Leu II	Ile Pro Leu 770 Val Pro	Gly Pro 755 Pro Gly Leu Leu	Thr 740 Glu Gly Ser Leu Glu 820	Val 725 Pro Pro Leu Ser Ala 805 Val	710 Pro Arg Pro Leu Ser 790 Ala Gly	Glu Asp Pro Ser 775 Gly Gln Leu	Glu Pro Thr 760 His Leu Gly Val	Glu Gly 745 His Gly Leu Gly Gly 825	Glu 730 Asp Leu Leu Pro Gly 810 Leu	715 Glu Gly Arg Leu 795 Gly	Glu Cys Pro Ala 780 Leu Leu	Glu Pro Cys 765 Gly Leu Gln Ser	Glu Ser 750 Pro Leu Leu Ala Tyr 830	Gly 735 Pro Ala Ser Leu Ala 815 Leu	Ala Asp Ser Phe Leu 800 Leu Leu
Ile II Gln II Ala V 785 Leu I	Ile Pro Leu 770 Val Pro	Gly Pro 755 Pro Gly Leu Leu Thr	Thr 740 Glu Gly Ser Leu Glu 820	Val 725 Pro Pro Leu Ser Ala 805 Val	710 Pro Arg Pro Leu Ser 790 Ala Gly	Glu Asp Pro Ser 775 Gly Gln Leu	Glu Pro Thr 760 His Leu Gly Val Pro	Glu Gly 745 His Gly Leu Gly Gly 825	Glu 730 Asp Leu Leu Pro Gly 810 Leu	715 Glu Gly Arg Leu 795 Gly	Glu Cys Pro Ala 780 Leu Leu	Glu Pro Cys 765 Gly Leu Gln Ser Leu	Glu Ser 750 Pro Leu Leu Ala Tyr 830	Gly 735 Pro Ala Ser Leu Ala 815 Leu	Ala Asp Ser Phe Leu 800 Leu Leu
Ile II Gln II 785 Leu II Leu II	Ile Pro Leu 770 Val Pro Ala	Gly Pro 755 Pro Gly Leu Leu Thr	Thr 740 Glu Gly Ser Leu Glu 820 Ala	Val 725 Pro Pro Leu Ser Ala 805 Val	710 Pro Arg Pro Leu Ser 790 Ala Gly His	Glu Asp Pro Ser 775 Gly Gln Leu Leu	Glu Pro Thr 760 His Leu Gly Val Pro 840	Glu Gly 745 His Gly Leu Gly Gly 825 Ser	Glu 730 Asp Leu Leu Pro Gly 810 Leu Ser	715 Glu Gly Arg Leu 795 Gly Gly Leu	Glu Cys Pro Ala 780 Leu Leu Ala Phe	Glu Pro Cys 765 Gly Leu Gln Ser Leu 845	Glu Ser 750 Pro Leu Leu Ala Tyr 830 Leu	Gly 735 Pro Ala Ser Leu Ala 815 Leu	Ala Asp Ser Phe Leu 800 Leu Leu Ala
Gln II	Ile Pro Leu 770 Val Pro Ala Cys	Gly Pro 755 Pro Gly Leu Leu Thr	Thr 740 Glu Gly Ser Leu Glu 820 Ala	Val 725 Pro Pro Leu Ser Ala 805 Val	710 Pro Arg Pro Leu Ser 790 Ala Gly His	Glu Asp Pro Ser 775 Gly Gln Leu Leu Ala	Glu Pro Thr 760 His Leu Gly Val Pro 840	Glu Gly 745 His Gly Leu Gly Gly 825 Ser	Glu 730 Asp Leu Leu Pro Gly 810 Leu Ser	715 Glu Gly Arg Leu 795 Gly Gly Leu	Glu Cys Pro Ala 780 Leu Leu Ala Phe Ser	Glu Pro Cys 765 Gly Leu Gln Ser Leu 845	Glu Ser 750 Pro Leu Leu Ala Tyr 830 Leu	Gly 735 Pro Ala Ser Leu Ala 815 Leu	Ala Asp Ser Phe Leu 800 Leu Leu Ala
Gln II	Ile Pro Leu 770 Val Pro Ala Cys Gly 850	Gly Pro 755 Pro Gly Leu Leu Thr 835 Thr	Thr 740 Glu Gly Ser Leu Glu 820 Ala	Val 725 Pro Pro Leu Ser Ala 805 Val Leu	710 Pro Arg Pro Leu Ser 790 Ala Gly His	Glu Asp Pro Ser 775 Gly Gln Leu Leu Ala 855	Glu Pro Thr 760 His Leu Gly Val Pro 840 Val	Glu Gly 745 His Gly Leu Gly 825 Ser Leu	Glu 730 Asp Leu Leu Pro Gly 810 Leu Ser	715 Glu Gly Arg Leu 795 Gly Gly Leu Leu	Glu Cys Pro Ala 780 Leu Leu Ala Phe Ser 860	Glu Pro Cys 765 Gly Leu Gln Ser Leu 845 Trp	Glu Ser 750 Pro Leu Leu Ala Tyr 830 Leu Arg	Gly 735 Pro Ala Ser Leu Ala 815 Leu Leu	Ala Asp Ser Phe Leu 800 Leu Leu Ala Gly

```
875
                    870
Pro Gly Leu Ala Leu Pro Leu Val Ala Met Ala Ala Gly Gly Arg Trp
                                   890
Val Arg Gln Gln Gly Pro Arg Val Arg Arg Gly Ile Ser Arg Leu Trp
                                905
                                                    910
            900
Leu Arg Val Leu Leu Arg Leu Ser Pro Met Ala Phe Arg Ala Leu Gln
                           920
        915
Gly Cys Gly Ala Val Gly Asp Arg Gly Leu Phe Ala Leu Tyr Pro Lys
                        935
                                            940
Thr Asn Lys Asp Gly Phe Arg Ser Arg Leu Pro Val Pro Gly Pro Arg
                    950
                                        955
Arg Arg Asn Pro Arg Thr Thr Gln His Pro Leu Ala Leu Leu Ala Arg
                                    970
                965
Val Trp Val Leu Cys Lys Gly Trp Asn Trp Arg Leu Ala Arg Ala Ser
                                985
Gln Gly Leu Ala Ser His Leu Pro Pro Trp Ala Ile His Thr Leu Ala
                            1000
        995
Ser Trp Gly Leu Leu Arg Gly Glu Arg Pro Thr Arg Ile Pro Arg Leu
                        1015
Leu Pro Arg Ser Gln Arg Gln Leu Gly Pro Pro Ala Ser Arg Gln Pro
                                        1035
                   1030
Leu Pro Gly Thr Leu Ala Gly Arg Arg Ser Arg Thr Arg Gln Ser Arg
                                    1050
Ala Leu Pro Pro Trp Arg
            1060
<210> 117
<211> 471
<212> DNA
<213> Homo sapiens
<400> 117
nacgcgttga cgatctgtct ggctggtgta gtgatctgcg ctgtgggtgt cgtcgatgac
ctgctcgacc ttcctgcctt ggccaaggca gctggccagg tattagcggc cggcatcgtc
gtcacgggcg gagtgcgaat gttttggatc ccgctgccga actccatcat tgctttgggg
acgcctactt cgatcttggt gacggtgttc ttcattgtgt tgtgcgccaa tgcggtgaat
ttcattgatg gacttgacgg cctggcatcc ggtgtggtgg ccatcgggtc cttggctttc
tteteataca cetacetget ggeteacgaa caggaetttg ttgttgcgae gaetaceagt
ctcattacgg ctgcgacggc gggcgcctgt ctcggttttt tgccccacaa ctggcatccg
gcgaggatgt tcatgggtga ttccggagct ctgctacttg gcttattgct a
<210> 118
<211> 157
<212> PRT
<213> Homo sapiens
```

```
<400> 118
Xaa Ala Leu Thr Ile Cys Leu Ala Gly Val Val Ile Cys Ala Val Gly
Val Val Asp Asp Leu Leu Asp Leu Pro Ala Leu Ala Lys Ala Ala Gly
                                25
Gln Val Leu Ala Ala Gly Ile Val Val Thr Gly Gly Val Arg Met Phe
                           40
Trp Ile Pro Leu Pro Asn Ser Ile Ile Ala Leu Gly Thr Pro Thr Ser
                       55
Ile Leu Val Thr Val Phe Phe Ile Val Leu Cys Ala Asn Ala Val Asn
                                       75
                   70
Phe Ile Asp Gly Leu Asp Gly Leu Ala Ser Gly Val Val Ala Ile Gly
                                   90
Ser Leu Ala Phe Phe Ser Tyr Thr Tyr Leu Leu Ala His Glu Gln Asp
                               105
Phe Val Val Ala Thr Thr Thr Ser Leu Ile Thr Ala Ala Thr Ala Gly
                            120
Ala Cys Leu Gly Phe Leu Pro His Asn Trp His Pro Ala Arg Met Phe
                        135
Met Gly Asp Ser Gly Ala Leu Leu Leu Gly Leu Leu
                    150
<210> 119
<211> 302
<212> DNA
<213> Homo sapiens
<400> 119
ntcaaacatg agcagtcgtg gcggccgagg ccgcggtggc tattatcgcg agctttatgg
tageegaggt egaggeagta aatetaatga aaetttegea aaaaattegg atgtetaete
tcagaaaaag actcgaacag tacgaggcac ctccgaagat ttagcacgat cgctccataa
getteatatg egecegtace etgegtatea tgacattgag ggtatgtggg ettteecage
ctttactttt tatctggatc atgcacaagc agacccatac gctgccccaa ataaggcacg
300
302
<210> 120
<211> 98
<212> PRT
<213> Homo sapiens
<400> 120
Met Ser Ser Arg Gly Gly Arg Gly Gly Tyr Tyr Arg Glu Leu
Tyr Gly Ser Arg Gly Arg Gly Ser Lys Ser Asn Glu Thr Phe Ala Lys
                                25
Asn Ser Asp Val Tyr Ser Gln Lys Lys Thr Arg Thr Val Arg Gly Thr
                            40
Ser Glu Asp Leu Ala Arg Ser Leu His Lys Leu His Met Arg Pro Tyr
```

```
60
                        55
Pro Ala Tyr His Asp Ile Glu Gly Met Trp Ala Phe Pro Ala Phe Thr
                                        75
                   70
Phe Tyr Leu Asp His Ala Gln Ala Asp Pro Tyr Ala Ala Pro Asn Lys
                                    90
                85
Ala Arg
<210> 121
<211> 318
<212> DNA
<213> Homo sapiens
<400> 121
ngcatggggg gccctgggac cgcacttgtg cccctttttt ttttagggaa aaaattgagc
cctaaaggat ttgccgcatt acaggaaagt tttttggtaa gtttggggtt gtttctgtgc
tgtgtgagaa ggagtagaag cagctccagt agagtgggcc ttttcatttt tatccagagg
aaatttgtag getgtggeta ttaetteett tttttettt tttttttttg tttagagaca
gagtetgnet etgtegeeag getggagtga agtggeaega teteagetea etgeaacete
tgcctcccag gttcaagc
318
<210> 122
<211> 89
<212> PRT
<213> Homo sapiens
<400> 122
Xaa Met Gly Gly Pro Gly Thr Ala Leu Val Pro Leu Phe Phe Leu Gly
                                    10
Lys Lys Leu Ser Pro Lys Gly Phe Ala Ala Leu Gln Glu Ser Phe Leu
Val Ser Leu Gly Leu Phe Leu Cys Cys Val Arg Arg Ser Arg Ser Ser
                            40
Ser Ser Arg Val Gly Leu Phe Ile Phe Ile Gln Arg Lys Phe Val Gly
                        55
Cys Gly Tyr Tyr Phe Leu Phe Phe Leu Phe Phe Phe Cys Leu Glu Thr
                                        75
                    70
Glu Ser Xaa Ser Val Ala Arg Leu Glu
                85
<210> 123
<211> 338
<212> DNA
<213> Homo sapiens
<400> 123
acgcgtctag ggtagaaatc aactccagta actgtcattc aacctcagca atgctgggga
```

```
cgggcagagg cagggcaget gtgtgccaca ttcctgccag ggctggtcag gccccggctc
teaccactee tecteeetge tttgaacetg tggaacaaag ggeeeetgea ceccaactea
tteetetttg ecacataagg geeteaagte atgetgteee etetgeetgg gttgettttt
ctccctctgc ttgggtcact gttcacacca ctggccactt tcctcaggga agggccctca
ctqcccacac acctaaacat gccccctgct cctccata
338
<210> 124
<211> 96
<212> PRT
<213> Homo sapiens
<400> 124
Met Leu Gly Thr Gly Arg Gly Arg Ala Ala Val Cys His Ile Pro Ala
                                    10
1
Arg Ala Gly Gln Ala Pro Ala Leu Thr Thr Pro Pro Pro Cys Phe Glu
                                25
                                                    30
            20
Pro Val Glu Gln Arg Ala Pro Ala Pro Gln Leu Ile Pro Leu Cys His
                            40
Ile Arg Ala Ser Ser His Ala Val Pro Ser Ala Trp Val Ala Phe Ser
Pro Ser Ala Trp Val Thr Val His Thr Thr Gly His Phe Pro Gln Gly
                                        75
                    70
Arg Ala Leu Thr Ala His Thr Pro Lys His Ala Pro Cys Ser Ser Ile
                                    90
                85
<210> 125
<211> 280
<212> DNA
<213> Homo sapiens
<400> 125
ccatggacct ggccagccac catcacctgc ctcctgcctc acccaccctg ggtgcctgcc
ggcaaggatt ggagggcaga ctgctggagc gtgagaccag gccaatctgt ctttctggga
accttcagcc tccaactgga gctgactgtc aactttcggg tgagaagtca cttttctgca
ttcccaccac actatctatc tgtgcaatac ggcagcgtga cagcactcac cttattgagg
gettetgetg teetggeeca ttetggatag geetgateta
<210> 126
<211> 92
<212> PRT
<213> Homo sapiens
<400> 126
Met Asp Leu Ala Ser His His Leu Pro Pro Ala Ser Pro Thr Leu
```

```
10
Gly Ala Cys Arg Gln Gly Leu Glu Gly Arg Leu Leu Glu Arg Glu Thr
                              25
Arg Pro Ile Cys Leu Ser Gly Asn Leu Gln Pro Pro Thr Gly Ala Asp
                                              45
                           40
Cys Gln Leu Ser Gly Glu Lys Ser Leu Phe Cys Ile Pro Thr Thr Leu
                                          60
                       55
Ser Ile Cys Ala Ile Arg Gln Arg Asp Ser Thr His Leu Ile Glu Gly
65
                   70
Phe Cys Cys Pro Gly Pro Phe Trp Ile Gly Leu Ile
<210> 127
<211> 444
<212> DNA
<213> Homo sapiens
<400> 127
cgcgtgatcg ccgtggcgga gggccgcggc gccgactcga tcgcccagct gacaaccgag
ctgcaaagcc gtcactgccc tgcggagcag atcacgtccg tcagcatcga catgtcgcca
gegttcatca ggggctgcgc cgagcacctg cccaacgcgc gcgtcacctt cgacaagttc
gacaagteee teaaggggat gegetggteg etgetgaaga acegegeeag eeteaageee
gaggetgeeg eegatetgga tgeeetgate geeaggatgg eeaetgtgeg eaeegegege
geetgggtet acaaggagea getgegegag ateetegege geaageagat caaegtggea
cgcgacatgc tcaagcactg gtgc
444
<210> 128
<211> 148
<212> PRT
<213> Homo sapiens
<400> 128
Arg Val Ile Ala Val Ala Glu Gly Arg Gly Ala Asp Ser Ile Ala Gln
Leu Thr Thr Glu Leu Gln Ser Arg His Cys Pro Ala Glu Gln Ile Thr
                               25
            20
Ser Val Ser Ile Asp Met Ser Pro Ala Phe Ile Arg Gly Cys Ala Glu
                                               45
                           40
His Leu Pro Asn Ala Arg Val Thr Phe Asp Lys Phe His Val Ile Gly
                                           60
                       55
His Ala Asn Ala Ala Val Asp Arg Met Arg Arg Ile Glu Gln Arg Ser
                                       75
                    70
Asp Lys Ser Leu Lys Gly Met Arg Trp Ser Leu Leu Lys Asn Arg Ala
                                   90
Ser Leu Lys Pro Glu Ala Ala Ala Asp Leu Asp Ala Leu Ile Ala Arg
```

```
100
                                105
Met Ala Thr Val Arg Thr Ala Arg Ala Trp Val Tyr Lys Glu Gln Leu
                            120
                                                125
Arg Glu Ile Leu Ala Arg Lys Gln Ile Asn Val Ala Arg Asp Met Leu
                        135
    130
Lys His Trp Cys
145
<210> 129
<211> 291
<212> DNA
<213> Homo sapiens
<400> 129
gaggagggac gtaccgtccc cgttatagcc aagctcgaga agccgcaagc tatcgagaac
ttggacgaga ttattgacgt ctttgacgcc gtcatggttg cccgtggcga tatggccgtc
qaqtqcccqc tcgaggaagt tccgctgatc caaaagcaga tcatcgagaa ggctcgttta
caqqctaaqc ccgtcattgt ggccacccag atgettgagt cgatgateca cgctccccgt
ccgacccgcg ctgaggccgc cgacgtcgcg aacgccatcc ttgacggcgc g
291
<210> 130
<211> 97
<212> PRT
<213> Homo sapiens
<400> 130
Glu Glu Gly Arg Thr Val Pro Val Ile Ala Lys Leu Glu Lys Pro Gln
Ala Ile Glu Asn Leu Asp Glu Ile Ile Asp Val Phe Asp Ala Val Met
                                25
Val Ala Arg Gly Asp Met Ala Val Glu Cys Pro Leu Glu Glu Val Pro
Leu Ile Gln Lys Gln Ile Ile Glu Lys Ala Arg Leu Gln Ala Lys Pro
                                            60
                        55
Val Ile Val Ala Thr Gln Met Leu Glu Ser Met Ile His Ala Pro Arg
                                        75
                    70
Pro Thr Arg Ala Glu Ala Ala Asp Val Ala Asn Ala Ile Leu Asp Gly
                                    90
Ala
<210> 131
<211> 416
<212> DNA
<213> Homo sapiens
<400> 131
teeggagegt eegtggeeet catgggtgtg teagegtggt tgetgteteg ggeegeagag
```

```
attecacegg tgetetacet ggaggeegea geegtegggg ttegattett eggeatetee
cgcggtgtct tccgctacgc cgaacgtctg gtaggccacg acctggctct gcggatgcag
180
ggggcattgc ggatgcgggt ctacgaccgg ctgtcacgta ccnaccctgc tgggnnacgt
egeeggggtg acetgetggt acgggttact geegaegteg acgeggtgtt ggacatggte
300
gtgegggtga tegtteegge gtgegegtea ageetegtea teattggeae eaeggteett
360
ctttgtccga gagaaggttg agttttctta gccggattcc aacacagcct gggggc
<210> 132
<211> 126
<212> PRT
<213> Homo sapiens
<400> 132
Ser Gly Ala Ser Val Ala Leu Met Gly Val Ser Ala Trp Leu Leu Ser
                                    1.0
Arg Ala Ala Glu Ile Pro Pro Val Leu Tyr Leu Glu Ala Ala Ala Val
            20
Gly Val Arg Phe Phe Gly Ile Ser Arg Gly Val Phe Arg Tyr Ala Glu
                            40
Arg Leu Val Gly His Asp Leu Ala Leu Arg Met Gln Gly Ala Leu Arg
                        55
Met Arg Val Tyr Asp Arg Leu Ser Arg Thr Xaa Pro Ala Gly Xaa Arg
                                        75
                    70
Arg Arg Gly Asp Leu Leu Val Arg Val Thr Ala Asp Val Asp Ala Val
                85
                                    90
Leu Asp Met Val Val Arg Val Ile Val Pro Ala Cys Ala Ser Ser Leu
                                105
            100
Val Ile Ile Gly Thr Thr Val Leu Leu Cys Pro Arg Glu Gly
                            120
        115
<210> 133
<211> 327
<212> DNA
<213> Homo sapiens
<400> 133
geegttgeta tegetgetgg tatgegtgea gaegteactg tttttgatat caatateget
gegttgaaga gactegeega catetaceag ggtegtgtte acaeagtagt atecaeeege
geogaaattg egaaggeget agaaaceget gaegttgtga teggttetgt eettatteeg
ggtagttcta ccccgaaget tgttactacc gatatggttg etcacatgca gcctgggtct
gttettattg atattgetat agaccaagge ggetgetteg aggattegea ecceaccaet
tacgatgacc ccactttcac tgtgcac
327
```

```
<210> 134
<211> 109
<212> PRT
<213> Homo sapiens
<400> 134
Ala Val Ala Ile Ala Ala Gly Met Arg Ala Asp Val Thr Val Phe Asp
                                    10
1
                 5
Ile Asn Ile Ala Ala Leu Lys Arg Leu Ala Asp Ile Tyr Gln Gly Arg
                                25
Val His Thr Val Val Ser Thr Arg Ala Glu Ile Ala Lys Ala Leu Glu
                                                 45
        35
Thr Ala Asp Val Val Ile Gly Ser Val Leu Ile Pro Gly Ser Ser Thr
                        55
Pro Lys Leu Val Thr Thr Asp Met Val Ala His Met Gln Pro Gly Ser
                    70
                                        75
Val Leu Ile Asp Ile Ala Ile Asp Gln Gly Gly Cys Phe Glu Asp Ser
                                    90
His Pro Thr Thr Tyr Asp Asp Pro Thr Phe Thr Val His
                                105
            100
<210> 135
<211> 560
<212> DNA
<213> Homo sapiens
<400> 135
taagatgtgg teetgeeetg tteetgaagg ggetgeaget etgatggaaa atacagggat
ttacactcag ggctacagcc acggggggct gaggcccaag gctgcaatct cgggggaagg
ggaagttggc ttttcctggt ggattggaaa catcctcttg gaggcaaaga cttttcctgg
180
atcttacaga cttcccggga tttttagatt agaatattgg gggcaaagga ggctgtcttg
ttttaaagca atgctacata gacacagtgg ggaagacctg gttcgacggc agataagcag
tgggtgatgg gcttgaggag gagagtcagg gcaaagtcta agactgagca gaaaggaatt
360
coccatcte ecatggataa gtacgtteta gaacattete tttgggteta atactetgaa
420
atgacatett gtetteatge tegagagaga attaetteae tggeteeaet tggagtgeea
gtgttcagac accaagcctg actgggaggg ttccgttttc ttaacacctt cccaccgccg
acttccaagt ccccacgcgt
560
<210> 136
 <211> 100
<212> PRT
<213> Homo sapiens
```

```
<400> 136
Met Trp Ser Cys Pro Val Pro Glu Gly Ala Ala Leu Met Glu Asn
                                    10
1
Thr Gly Ile Tyr Thr Gln Gly Tyr Ser His Gly Gly Leu Arg Pro Lys
                                25
            20
Ala Ala Ile Ser Gly Glu Gly Glu Val Gly Phe Ser Trp Trp Ile Gly
                                                45
                            40
Asn Ile Leu Leu Glu Ala Lys Thr Phe Pro Gly Ser Tyr Arg Leu Pro
Gly Ile Phe Arg Leu Glu Tyr Trp Gly Gln Arg Arg Leu Ser Cys Phe
                    70
Lys Ala Met Leu His Arg His Ser Gly Glu Asp Leu Val Arg Arg Gln
                                    90
Ile Ser Ser Gly
            100
<210> 137
<211> 429
<212> DNA
<213> Homo sapiens
<400> 137
accggttgga tggcctgcag gccaaagcgt tcctgcaaac tcagcaggcc ttcagcgcaa
gaggcaaaca gctggtcgcg cacctgcttg aggtccaccg attgcgcatc gcccttgagc
aaggegegee agttggtttt gteggeeact tggetgegga acaggtette gacaaaaceg
gactgctggc gggtcgcaac gcgcatgatc ggcagcgcct ggctggcgcc ctggtcgagc
cagegegteg geagttgggt ggeeegggtg atacegaeet tgateeeega egaattggee
aggtacacca catggtcggt catgcagaat gtttcgcccc agccgggatc acggcaagtg
ccggcgtcgt aatggcaacg ttcggggctc atgatgcaca ggtcacactg ggccagcttg
420
gtcatgccc
429
<210> 138
<211> 141
<212> PRT
<213> Homo sapiens
<400> 138
Met Thr Lys Leu Ala Gln Cys Asp Leu Cys Ile Met Ser Pro Glu Arg
                                    10
Cys His Tyr Asp Ala Gly Thr Cys Arg Asp Pro Gly Trp Gly Glu Thr
                                25
Phe Cys Met Thr Asp His Val Val Tyr Leu Ala Asn Ser Ser Gly Ile
        35
Lys Val Gly Ile Thr Arg Ala Thr Gln Leu Pro Thr Arg Trp Leu Asp
                                            60
                        55
Gln Gly Ala Ser Gln Ala Leu Pro Ile Met Arg Val Ala Thr Arg Gln
```

```
75
                   70
Gln Ser Gly Phe Val Glu Asp Leu Phe Arg Ser Gln Val Ala Asp Lys
                                   90
Thr Asn Trp Arg Ala Leu Leu Lys Gly Asp Ala Gln Ser Val Asp Leu
                               105
            100
Lys Gln Val Arg Asp Gln Leu Phe Ala Ser Cys Ala Glu Gly Leu Leu
                           120
       115
Ser Leu Gln Glu Arg Phe Gly Leu Gln Ala Ile Gln Pro
                       135
<210> 139
<211> 341
<212> DNA
<213> Homo sapiens
<400> 139
acgcgtcgtt tgaaggcttg atccgcacgt ccaattcgct ttgcgccaat gcgccgcagc
ttgtgaacag cagaatcaag ccgctggtaa atcttcctgg gagcttcata ggcggggatg
ctacacgage tggggagaca ctttgaacce ggaattgtet gaataattet gteteaaace
tttgcagcct gtaacgactg agggttcgga tggaaaaaca catgctccag gatgggaccg
acggccactt caccgatete tteatagece tggcgtttgt agaaatecag gtagegegaa
togocagogt cgagcacgac gcctgatgag tgcgggtcat t
341
<210> 140
<211> 113
<212> PRT
<213> Homo sapiens
<400> 140
Met Thr Arg Thr His Gln Ala Ser Cys Ser Thr Leu Ala Ile Arg Ala
Thr Trp Ile Ser Thr Asn Ala Arg Ala Met Lys Arg Ser Val Lys Trp
                                25
            20
Pro Ser Val Pro Ser Trp Ser Met Cys Phe Ser Ile Arg Thr Leu Ser
                            40
Arg Tyr Arg Leu Gln Arg Phe Glu Thr Glu Leu Phe Arg Gln Phe Arg
                        55
Val Gln Ser Val Ser Pro Ala Arg Val Ala Ser Pro Pro Met Lys Leu
                                        75
                    70
Pro Gly Arg Phe Thr Ser Gly Leu Ile Leu Leu Phe Thr Ser Cys Gly
                                    90
Ala Leu Ala Gln Ser Glú Leu Asp Val Arg Ile Lys Pro Ser Asn Asp
                                105
            100
Ala
<210> 141
<211> 324
```

```
<212> DNA
<213> Homo sapiens
<400> 141
gaatteetet tggatagett egggtaaatg ggtaeageaa atateaggag egeaacegea
acctttactt actggtacat gaacaccatt tacattacag ctatcgtact caccecacgt
catgtgaaca gacacataac tgaaaggttt ataaaccaca gtctcacggt acgtatgacc
gtcaactgtg aacaccgcta agtaatagcc tgcgggggct tgcatgaact cctttgacca
tgcgtaataa atacgtccgt cattagtcac acctgatggg gcgaaacaaa aagaacggca
gcagttatca ccgcccatac gcgt
324
<210> 142
<211> 106
<212> PRT
<213> Homo sapiens
<400> 142
Met Gly Gly Asp Asn Cys Cys Arg Ser Phe Cys Phe Ala Pro Ser Gly
                                    10
Val Thr Asn Asp Gly Arg Ile Tyr Tyr Ala Trp Ser Lys Glu Phe Met
                                25
            20
Gln Ala Pro Ala Gly Tyr Tyr Leu Ala Val Phe Thr Val Asp Gly His
                            40
Thr Tyr Arg Glu Thr Val Val Tyr Lys Pro Phe Ser Tyr Val Ser Val
                        55
His Met Thr Trp Gly Glu Tyr Asp Ser Cys Asn Val Asn Gly Val His
                                        75
                    70
Val Pro Val Ser Lys Gly Cys Gly Cys Ala Pro Asp Ile Cys Cys Thr
                85
                                     90
His Leu Pro Glu Ala Ile Gln Glu Glu Phe
            100
<210> 143
<211> 1325
<212> DNA
<213> Homo sapiens
<400> 143
nacgcgtgga tctgccagct gagcctggag ctgtgcaggc agctgccctg ctacgatgag
gcaccccagg agaagaactt cctgtacaaa tgcataggca ccaccctggg tgctgcttca
agtaaggagg tggtgaggaa gcaccttcaa gagctgctgg agacggccag ataccaggag
gaggeagaac gegagggeet egeetgetge ttegggatet gtgeeatete eeacetegag
gacacgetgg cecagetgga ggacttegtg aggteagagg tetteagaaa atecattgge
300
```

```
attctcaaca tttttaagga tcgaagtgag aacgaagtgg agaaggtgaa gagtgctctg
atcotgtgct atgggcacgt ggcggcccgg gcccccggg agctggtgct ggccaaggta
gagtcagaca tcctccggaa catcntgcca gcacttcagc acnncaagga cccagccctg
aagetgtgee ttgteeagag tgtgtgeatg gteageegeg eeatetgeag eageaceeag
getggeteet tecaetteae eeggaaagea gagetggtgg cacagatgat ggagtteate
agggcagage ecceggaete ettgaggaea ectattegga agaaageeat geteaeetge
acttacttgg tctccgtgga gccagcgctg gacgagcagg cccgggcgga tgtgatccat
ggetgeetge acageateat ggeeetgetg cetgageeea aggaggagga eggaggetge
cagaagteee tgtatetgga gacaetgeae gecettgagg atetgetgae gageeteetg
cageggaaca tgacececa aggeetgeag ateatgattg ageacetgag eccatggate
aagteeceaa gaggteaegt ageggegegt geectaggee tgagegeeet cetegtgege
960
tacttcctgg agcacctgcg tgtcagtggc gcccaagtag ataccaggtt tccatctgag
cccaggatcc tgtgcaatgg ccctggtgcc cttccacaac ctgggccttc tcatcggcct
cttctcccca cggtgtgcgg acctgtggcc tgccacccgc caggaggccg tggactgtgt
1140
ctactccctg ctgtacctcc agctcggcta tgagggcttc tcccgggact accgcgatga
cgtggcggag cggctcctca gcctcaagga cggcctcgtg caccctgacc ccgccattct
cttccacace tgccacagtg taggccagat tattgccaag cgcctcccc cagcccttca
1320
cgcgt
1325
<210> 144
<211> 390
<212> PRT
<213> Homo sapiens
<400> 144
Xaa Ala Trp Ile Cys Gln Leu Ser Leu Glu Leu Cys Arg Gln Leu Pro
Cys Tyr Asp Glu Ala Pro Gln Glu Lys Asn Phe Leu Tyr Lys Cys Ile
                                25
Gly Thr Thr Leu Gly Ala Ala Ser Ser Lys Glu Val Val Arg Lys His
        35
                            40
Leu Gln Glu Leu Leu Glu Thr Ala Arg Tyr Gln Glu Glu Ala Glu Arg
                                            60
Glu Gly Leu Ala Cys Cys Phe Gly Ile Cys Ala Ile Ser His Leu Glu
                    70
Asp Thr Leu Ala Gln Leu Glu Asp Phe Val Arg Ser Glu Val Phe Arg
```

```
Lys Ser Ile Gly Ile Leu Asn Ile Phe Lys Asp Arg Ser Glu Asn Glu
                105 110
         100
Val Glu Lys Val Lys Ser Ala Leu Ile Leu Cys Tyr Gly His Val Ala
                      120
                                        125
      115
Ala Arg Ala Pro Arg Glu Leu Val Leu Ala Lys Val Glu Ser Asp Ile
                   135
Leu Arg Asn Ile Xaa Pro Ala Leu Gln His Xaa Lys Asp Pro Ala Leu
    150 155
Lys Leu Cys Leu Val Gln Ser Val Cys Met Val Ser Arg Ala Ile Cys
            165 170
Ser Ser Thr Gln Ala Gly Ser Phe His Phe Thr Arg Lys Ala Glu Leu
                          185 190
         180
Val Ala Gln Met Met Glu Phe Ile Arg Ala Glu Pro Pro Asp Ser Leu
                      200
Arg Thr Pro Ile Arg Lys Lys Ala Met Leu Thr Cys Thr Tyr Leu Val
                   215
                                    220
Ser Val Glu Pro Ala Leu Asp Glu Gln Ala Arg Ala Asp Val Ile His
               230
                               235
Gly Cys Leu His Ser Ile Met Ala Leu Leu Pro Glu Pro Lys Glu Glu
                            250
            245
Asp Gly Gly Cys Gln Lys Ser Leu Tyr Leu Glu Thr Leu His Ala Leu
         260 265
                                          270
Glu Asp Leu Leu Thr Ser Leu Leu Gln Arg Asn Met Thr Pro Gln Gly
                                        285
                      280
Leu Gln Ile Met Ile Glu His Leu Ser Pro Trp Ile Lys Ser Pro Arg
                   295
                                    300
Gly His Val Ala Ala Arg Ala Leu Gly Leu Ser Ala Leu Leu Val Arg
                310
                                 315
Tyr Phe Leu Glu His Leu Arg Val Ser Gly Ala Gln Val Asp Thr Arg
             325
                              330
Phe Pro Ser Glu Pro Arg Ile Leu Cys Asn Gly Pro Gly Ala Leu Pro
                          345
          340
Gln Pro Gly Pro Ser His Arg Pro Leu Leu Pro Thr Val Cys Gly Pro
    355 360
Val Ala Cys His Pro Pro Gly Gly Arg Gly Leu Cys Leu Leu Pro Ala
  370 375
Val Pro Pro Ala Arg Leu
                390
<210> 145
<211> 802
<212> DNA
<213> Homo sapiens
<400> 145
eggeegteta ggteeggete agtgegetgt tgetegeegt agaacacgag getgegeaag
cataagcaga cgtagagagt ggtcacatcc atgtcgatgg tgtgcgcgta atgaaggtct
acatcaccct ggtgaaggcc tgcaccacta gcgtcggcac catttccccg cgtcggacaa
gacatcatge eccatatett gacagaatgt etgacatgag tatgecaege egageageae
240
```

```
cagaggacga caccgatetg geggacgeeg ecegtteatg gegeagatae eteateeteg
teatttgtgg egttategte getgteeteg gaetaggeat tttegggtat ettgegtggt
360
ggtcattgtg cgatcaagct gccggggtct gtcagcgtgg tgaacccgtt atgtactggt
420
gtteggtggt etetetggee atteteggae teattategg ggtettgaeg eagatetgge
tggagaagcg ctggtggcac atgcttgcca tcgtcatccc ggctgttttc atcgtcgccg
gtatettttt etggetegee gtetaagaag gggegteaca gattecacaa aegacacagg
600
tattgatete egitttateg geteetagea geegiggiea aegiateget ateaagegat
acaggactcg tcgttcgcat cgttgttgtg ctgctgggaa acaatcccag cgatctactc
ggetacegee agacagttea etcacaacee etcacgeegg egcagacate aaateeeatt
ctcgatagac ggcccacacc ac
802
<210> 146
<211> 151
<212> PRT
<213> Homo sapiens
<400> 146
Met Lys Val Tyr Ile Thr Leu Val Lys Ala Cys Thr Thr Ser Val Gly
Thr Ile Ser Pro Arg Arg Thr Arg His His Ala Pro Tyr Leu Asp Arg
                                25
            20
Met Ser Asp Met Ser Met Pro Arg Arg Ala Pro Glu Asp Asp Thr
                                                 45
                            40
Asp Leu Ala Asp Ala Ala Arg Ser Trp Arg Arg Tyr Leu Ile Leu Val
                        55
                                            60
Ile Cys Gly Val Ile Val Ala Val Leu Gly Leu Gly Ile Phe Gly Tyr
                    70
Leu Ala Trp Trp Ser Leu Cys Asp Gln Ala Ala Gly Val Cys Gln Arg
                                    90
                85
Gly Glu Pro Val Met Tyr Trp Cys Ser Val Val Ser Leu Ala Ile Leu
                                105
Gly Leu Ile Ile Gly Val Leu Thr Gln Ile Trp Leu Glu Lys Arg Trp
                                                 125
                            120
        115
Trp His Met Leu Ala Ile Val Ile Pro Ala Val Phe Ile Val Ala Gly
                        135
Ile Phe Phe Trp Leu Ala Val
                    150
145
<210> 147
<211> 368
<212> DNA
<213> Homo sapiens
<400> 147
```

```
acgcgtgaaa acggtatgac tettetggee ttagtagate tgtetaaaaa accegatgag
tttacacagt qqqcattaqt aqcccqcgat gttcatgaca ttcctggtct acgaaaagtt
attggtcaga aagtaccttg tgttgcagtg acggggtcgg aaaaggtgct tcataaaaag
gattactggg atctagcaac acctatgcca attgcgtggg gtacaacgga ccgaacagtt
attgctgatg cacqacqtac aatccccacc acggagtggg atatccttgc aagactacgt
ccacgcctag aagaggttcg caagcaacgt aatgatgtat tgctcctcaa cgaggaggat
360
cccccta
368
<210> 148
<211> 117
<212> PRT
<213> Homo sapiens
<400> 148
Met Thr Leu Leu Ala Leu Val Asp Leu Ser Lys Lys Pro Asp Glu Phe
Thr Gln Trp Ala Leu Val Ala Arg Asp Val His Asp Ile Pro Gly Leu
            20
Arg Lys Val Ile Gly Gln Lys Val Pro Cys Val Ala Val Thr Gly Ser
                            40
Glu Lys Val Leu His Lys Lys Asp Tyr Trp Asp Leu Ala Thr Pro Met
                                            60
                        55
Pro Ile Ala Trp Gly Thr Thr Asp Arg Thr Val Ile Ala Asp Ala Arg
                    70
                                        75
Arg Thr Ile Pro Thr Thr Glu Trp Asp Ile Leu Ala Arg Leu Arg Pro
                85
                                    90
Arg Leu Glu Glu Val Arg Lys Gln Arg Asn Asp Val Leu Leu Leu Asn
                                105
            100
Glu Glu Asp Pro Pro
        115
<210> 149
<211> 407
<212> DNA
<213> Homo sapiens
<400> 149
nngctagcat ggaccctagt cacacaggca gccatacccg aggtcaaagt gacccatttt
cctaatatgg ccgctcagat ccaatacttt gaagattcgt ccgtggttat atggcacgat
geggtggatg gtategtgta eegaagtgeg gatgaaggea agtegtggge eecaattaag
gggcctgaac agggtcaggc gcaccttttc gtgctccatc cctacgacaa gactcaagcg
tatattetga egegeageae teageattgg egeaegtega acegtggega gaegtggeag
300
```

```
tcattctcaa cycctcatcc ycctacgacc ttgaaagcta tycctctgga ctttcatccg
acgcatcatg actggatcct tttcacgggc caggcttgca cggtaaa
407
<210> 150
<211> 135
<212> PRT
<213> Homo sapiens
<400> 150
Xaa Leu Ala Trp Thr Leu Val Thr Gln Ala Ala Ile Pro Glu Val Lys
Val Thr His Phe Pro Asn Met Ala Ala Gln Ile Gln Tyr Phe Glu Asp
Ser Ser Val Val Ile Trp His Asp Ala Val Asp Gly Ile Val Tyr Arg
Ser Ala Asp Glu Gly Lys Ser Trp Ala Pro Ile Lys Gly Pro Glu Gln
Gly Gln Ala His Leu Phe Val Leu His Pro Tyr Asp Lys Thr Gln Ala
                    70
Tyr Ile Leu Thr Arg Ser Thr Gln His Trp Arg Thr Ser Asn Arg Gly
                                    90
                85
Glu Thr Trp Gln Ser Phe Ser Thr Pro His Pro Pro Thr Thr Leu Lys
            100
                                105
                                                    110
Ala Met Pro Leu Asp Phe His Pro Thr His His Asp Trp Ile Leu Phe
                            120
                                                125
       115
Thr Gly Gln Ala Cys Thr Val
   130
                        135
<210> 151
<211> 448
<212> DNA
<213> Homo sapiens
<400> 151
accggtgtcc gtggctattg ccccgaatgg tccccatccg cgtccccggg aactccctcg
gettttegeg catecaggte eccageecca getaetggtg egeeeegage ecctaggtge
cagagoggtg gtoggooggg ctootgooca gtotoggoto otocotooto cocaccagaa
ggaaaaactt gggcccttcg agaaccctgt ggaatgttct ttgtaatcaa ctgtacatcc
gettecacgg caeggeeteg tgcaaaateg egggtttegg ggeettggag caaattgege
300
ttgtcagcgg cgacgtcagg aggacaaggg gaggggttcg cggctgaaac tgcagcttcg
cagcacagag ccattttagg ctgctcccca cctcgcgggg cccatgggaa gccggccccg
ggagggcgcg gctgcatgga tattcgac
448
```

<210> 152

```
<211> 149
<212> PRT
<213> Homo sapiens
<400> 152
Thr Gly Val Arg Gly Tyr Cys Pro Glu Trp Ser Pro Ser Ala Ser Pro
Gly Thr Pro Ser Ala Phe Arg Ala Ser Arg Ser Pro Ala Pro Ala Thr
                                25
            20
Gly Ala Pro Arg Ala Pro Arg Cys Gln Ser Gly Gly Arg Pro Gly Ser
                                                 45
Cys Pro Val Ser Ala Pro Pro Ser Ser Pro Pro Glu Gly Lys Thr Trp
                        55
Ala Leu Arg Glu Pro Cys Gly Met Phe Phe Val Ile Asn Cys Thr Ser
                                        75
65
Ala Ser Thr Ala Arg Pro Arg Ala Lys Ser Arg Val Ser Gly Pro Trp
                                    90
                85
Ser Lys Leu Arg Leu Ser Ala Ala Thr Ser Gly Gly Gln Gly Glu Gly
            100
Phe Ala Ala Glu Thr Ala Ala Ser Gln His Arg Ala Ile Leu Gly Cys
                            120
Ser Pro Pro Arg Gly Ala His Gly Lys Pro Ala Pro Gly Gly Arg Gly
                        135
Cys Met Asp Ile Arg
145
<210> 153
<211> 440
<212> DNA
<213> Homo sapiens
<400> 153
nnntgggtcc atgtatgtgt gtgtatatga gggagacacg caggtgtgtg tccgagtgtg
tgtccatggg tccatgtatg tgtgtgtata tgtggggggaa caggtgtgtg tccgagtgtg
tgcatgggtc cgtgtatatg cgtgtatata tgcggggata tgtatatgtg tgtgtgtatg
aacaggtgta agtggggagc actcaggtgt gtctgtgtgt gttcgtgtac acgtgtgtaa
gtgggtgacc atgaaggggt gtgtgtgtcc gtgtgtaggt ttgcgtgcat gcacacatgc
atgtgtgtac tggggcatcc aagcccctgg tetecactcc attecaccct acgcctacct
360
cettgatete tgegeccage ettggetgtg eteceetget gtatgeaegt gggtgtetge
420
acgtgggtgt ctgcacgcgt
440
<210> 154
<211> 69
<212> PRT
<213> Homo sapiens
```

```
<400> 154
Gly Arg His Ala Gly Val Cys Pro Ser Val Cys Pro Trp Val His Val
                                    10
Cys Val Cys Ile Cys Gly Gly Thr Gly Val Cys Pro Ser Val Cys Met
            20
Gly Pro Cys Ile Cys Val Tyr Ile Cys Gly Asp Met Tyr Met Cys Val
                            40
Cys Met Asn Arg Cys Lys Trp Gly Ala Leu Arg Cys Val Cys Val Cys
                        55
Ser Cys Thr Arg Val
65
<210> 155
<211> 344
<212> DNA
<213> Homo sapiens
<400> 155
acgcgtatcg accaccatgt cgtcgtcacc acggcaagcg ctctcggcgg gcgagaacga
gtgaacatgg ccgagttgat ggccgatgcc gcgaccggca cgaaaccgtc ctacctacag
egatetteet eetegateae etegtttgaa gtggacaggg aacaaagaca eteagacaae
gegeegeagg aagtaaaaag ttegetetee gateaeggee gtegegegag tgeaeaggga
gaactgggca cotogcaago tacgccaccg cgatccatgo coccgcccgt atottccgcc
tectetacet ecceettace gateageatt atateegate taga
344
<210> 156
<211> 92
<212> PRT
<213> Homo sapiens
<400> 156
Met Ala Glu Leu Met Ala Asp Ala Ala Thr Gly Thr Lys Pro Ser Tyr
                                    10
Leu Gln Arg Ser Ser Ser Ile Thr Ser Phe Glu Val Asp Arg Glu
            20
Gln Arg His Ser Asp Asn Ala Pro Gln Glu Val Lys Ser Ser Leu Ser
                             40
Asp His Gly Arg Arg Ala Ser Ala Gln Gly Glu Leu Gly Thr Ser Gln
                                             60
                         55
Ala Thr Pro Pro Arg Ser Met Pro Pro Pro Val Ser Ser Ala Ser Ser
                     70
Thr Ser Pro Leu Pro Ile Ser Ile Ile Ser Asp Leu
                85
<210> 157
 <211> 6816
 <212> DNA
 <213> Homo sapiens
```

<400> 157					
nnagatetee	aaagaaccaa	aaggatcaat	atttctggat	teetgtatgg	gtgtcgttca
gaacaacaaa 120	gtcaggcgtt	ttgcttttga	gctcaagatg	caggacaaaa	gtagttatct
cttggcagca	gacagtgaag	tggaaatgga	agaatggatc	acaattctaa	ataagatcct
ccagctcaac		caatgcaaga			
tgaacaaagc 300		gttctggttc			
360		aaatcaaact			
420		agcttgactt			
480		ggateettgt			
540		aagaaggacc			
600		aatacaaccg			
660		aaatgatcgc			
720		cctcaggggc			
780		cacttgtcct			
840		catcacacat			
900					aagactagga
960					atctggaaat
1020					gctatccaat
1080					taagctccca
1140					ccctaattat
1200					tcccatcacg
1260					caccatctac
1320					gtettttgee
1380					. agacteteag
1440					cgcctttgct
1500					. agagttgccc
actcagctgo 1560	: atgaaaagca	. ccacctgttg	ctcacattçt	tccatgtcag	ctgtgacaac

1620			gtcgttgaaa		
cttcccctcc 1680	tgaaagacgg	aagggtggtg	acaagcgagc	agcacatccc	ggtctcggcg
aaccttcctt 1740	cgggctatct	tggctaccag	gagcttggga	tgggcaggca	ttatggtccg
gaaattaaat 1800	gggtagatgg	aggcaagcca	ctgctgaaaa	tttccactca	tctggtttct
acagtgtata 1860	ctcaggatca	gcatttacat	aatttttcc	agtactgtca	gaaaaccgaa
tctggagece 1920	aagccttagg	aaacgaactt	gtaaagtacc	ttaagagtct	gcatgcgatg
gaaggccacg 1980	tgatgatcgc	cttcttgccc	actatcctaa	accagctgtt	ccgagtcctc
accagagcca 2040	cacaggaaga	agtcgcggtt	aacgtgactc	gggtcattat	tcatgtggtt
gcccagtgcc 2100	atgaggaagg	attggagagc	cacttgaggt	catatgttaa	gtacgcgtat
aaggetgage 2160	catatgttgc	ctctgaatac	aagacagtgc	atgaagaact	gaccaaatcc
atgaccacga 2220	ttctcaagcc	ttctgccgat	ttcctcacca	gcaacaaact	actgaagtac
tcatggtttt 2280	tctttgatgt	actgatcaaa	tetatggete	agcatttgat	agagaactcc
aaagttaagt 2340	tgctgcgaaa	ccagagattt	cctgcatcct	atcatcatgc	agtggaaacc
gttgtaaata 2400	tgctgatgcc	acacatcact	cagaagtttc	gagataatcc	agaggcatct
aagaacgcga 2460	atcatagcct	tgetgtette	atcaagagat	gtttcacctt	catggacagg
ggctttgtct 2520	tcaagcagat	caacaactac	attagctgtt	ttgctcctgg	agacccaaag
accetetttg 2580	aatacaagtt	tgaatttctc	cgtgtagtgt	gcaaccatga	acattatatt
ccgttgaact 2640	taccaatgcc	atttggaaaa	ggcaggattc	aaagatacca	agacctccag
cttgactact 2700	cattaacaga	tgagttctgc	agaaaccact	tcttggtggg	actgttactg
agggaggtgg 2760	ggacagccct	ccaggagttc	cgggaggtcc	gtctgatcgc	catcagtgtg
ctcaagaacc 2820	tgctgataaa	gcattcttt	gatgacagat	atgcttcaag	gagccatgag
gcaaggatag 2880	ccacceteta	catgactatg	tttggtctgc	tgattgaaaa	cgtccagcgg
atcaatgtga 2940	gggatgtgtc	accettecet	gtgaacgcgg	gcatgactgt	gaaggatgaa
teeetggete 3000	taccagetgt	gaateegetg	gtgacgccgc	agaagggaag	caccctggac
aacagcctgc 3060	acaaggacct	gctgggcgcc	atctccggca	ttgcttctcc	atatacaacc
	acatcaacag	tgtgagaaat	gctgattcga	gaggatetet	cataagcaca
	acageettee	agaaaggaat	agtgagaaga	gcaattccct	ggataagcac

caacaaagta 3240	gcacattggg	aaattccgtg	gttcgctgtg	ataaacttga	ccagtctgag
attaagagcc 3300	tactgatgtg	tttcctctac	atcttaaaga	gcatgtctga	tgatgctttg
tttacatatt 3360	ggaacaaggc	ttcaacatct	gaacttatgg	attttttac	aatatctgaa
gtctgcctgc 3420	accagttcca	gtacatgggg	aagcgataca	tagccagaac	aggaatgatg
catgccagat 3480	tgcagcagct	gggcagcctg	gataactctc	tcacttttaa	ccacagctat
ggccactcgg 3540	acgcagatgt	tctgcaccag	tcattacttg	aagccaacat	tgctactgag
gtttgcctga 3600	cagctctgga	cacgctttct	ctatttacat	tggcgtttaa	gaaccagctc
ctggccgacc 3660	atggacataa	tecteteatg	aaaaaagttt	ttgatgtcta	cctgtgtttt
3720			_	tcactgcctt	
3780				acatgtgtgc	
3840				tcaggacgga	
3900		_	_	gaaagaagtc	
3960				cagacgttgt	
4020		_		gtgccaacag	
4080				ccaaaaggat	
4140				cagagatgct	
4200				tcaggaagac	
4260				aggcagcaat	
4320		_		aagcagtcca	
4380	_			gccggggagg	
4440				acgaggaggc	
4500				tgctgatgga	•
4560				tcattgccga	
cttatcatcc 4620	ccatttatga	gaagcggagg	gattttgaga	ggctggccca	tctgtatgac
acgetgeace 4680	gggcctacag	caaagtgacc	gaggtcatgc	actcgggccg	caggettetg
4740				aataccagtt	
gaaacagatg 4800	tggagggatt	ctttgaagat	gaagatggaa	aggagtatat	ttacaaggaa

cccaaactca 4860	caccgctgtc	ggaaatttct	cagagactcc	ttaaactgta	ctcggataaa
tttggttctg 4920	aaaatgtcaa	aatgatacag	gattetggca	aggtcaaccc	taaggatctg
gattctaagt 4980	atgcctacat	ccaggtgact	cacgtcatcc	ccttctttga	cgaaaaagag
ttgcaagaaa 5040	ggaaaacaga	gtttgagaga	tcccacaaca	tccgccgctt	catgtttgag
atgccattta 5100	cgcagaccgg	gaagaggcag	ggcggggtgg	aagagcagtg	caaacggcgc
5160	cagecataca				
cagcaccaca 5220	ctgacctgaa	ccccatcgag	gtggccattg	acgagatgag	taagaaggtg
5280	ggcagctgtg				
ctccagggca 5340	gcgtgagtgt	tcaggtcaat	getggeecae	tagcatatgc	gcgagctttc
5400	caaacacaaa				
5460	ttgtggaagc				
gaagaccagc 5520	tcgagtatca	ggaagaaatg	aaagccaact	acagggaaat	ggcgaaggag
5580	tcatgcatga				
5640	acatcttcaa				
5700	cgtcttcggt				
5760	caaactcagg	_	_		
5820	gggaagggga				
5880	gttgtaagaa				
5940	tgtcttaaca				
6000	agagtagtgt				
6 06 0	gagggctgca				
6120	cttgaaaata				
6180	gactagtttt				
gtgaaccagg 6240	agatttagtg	cttttatatt	catttccttg	catttaagaa	aatatgaaag
6300	tatgtgaget				
6360	ctatgctgaa				
tgttaataca 6420	gtcctggcat	tctgtacata	tatgtattac	atttctacat	ttttaatact

```
cacatgggct tatgcattaa gtttaattgt gataaatttg tgctgttcca gtatatgcaa
6480
tacactttaa tgttttattc ttgtacataa aaatgtgcaa tatggagatg tatacagtct
ttactatatt aggtttataa acagttttaa gaatttcatc cttttgccaa aatggtggag
tatgtaattg gtaaatcata aatcctgtgg tgaatggtgg tgtactttaa agctgtcacc
atgttatatt ttcttttaag acattaattt agtaatttta tatttgggaa aataaaggtt
tttaatttta tttaactgga atcactgccc tgctgtaatt aaacattctg taccacatct
gtattaaaaa gacattgctg accattaaaa aaaaaa
6816
<210> 158
<211> 1572
<212> PRT
<213> Homo sapiens
<400> 158
Ala Ser Gly Asn Leu Asp Lys Asn Ala Arg Phe Ser Ala Ile Tyr Arg
Gln Asp Ser Asn Lys Leu Ser Asn Asp Asp Met Leu Lys Leu Leu Ala
                                25
Asp Phe Arg Lys Pro Glu Lys Met Ala Lys Leu Pro Val Ile Leu Gly
                            40
Asn Leu Asp Ile Thr Ile Asp Asn Val Ser Ser Asp Phe Pro Asn Tyr
                        55
Val Asn Ser Ser Tyr Ile Pro Thr Lys Gln Phe Glu Thr Cys Ser Lys
                                         75
Thr Pro Ile Thr Phe Glu Val Glu Glu Phe Val Pro Cys Ile Pro Lys
                                    90
                85
His Thr Gln Pro Tyr Thr Ile Tyr Thr Asn His Leu Tyr Val Tyr Pro
                                105
Lys Tyr Leu Lys Tyr Asp Ser Gln Lys Ser Phe Ala Lys Ala Arg Asn
 Ile Ala Ile Cys Ile Glu Phe Lys Asp Ser Asp Glu Glu Asp Ser Gln
                                            140
Pro Leu Lys Cys Ile Tyr Gly Arg Pro Gly Gly Pro Val Phe Thr Arg
                                         155
                     150
Ser Ala Phe Ala Ala Val Leu His His His Gln Asn Pro Glu Phe Tyr
                                     170
Asp Glu Ile Lys Ile Glu Leu Pro Thr Gln Leu His Glu Lys His His
                                 185
            180
Leu Leu Leu Thr Phe Phe His Val Ser Cys Asp Asn Ser Ser Lys Gly
                             200
 Ser Thr Lys Lys Arg Asp Val Val Glu Thr Gln Val Gly Tyr Ser Trp
                                             220
                         215
Leu Pro Leu Leu Lys Asp Gly Arg Val Val Thr Ser Glu Gln His Ile
                                         235
                    230
 Pro Val Ser Ala Asn Leu Pro Ser Gly Tyr Leu Gly Tyr Gln Glu Leu
                                     250
                 245
 Gly Met Gly Arg His Tyr Gly Pro Glu Ile Lys Trp Val Asp Gly Gly
```

			260					265					270		
Lys	Pro	Leu 275		Lys	Ile	Ser	Thr 280		Leu	Val	Ser	Thr 285	Val	Tyr	Thr
Gln	Asp 290	Gln	His	Leu	His	Asn 295	Phe	Phe	Gln	Tyr	Cys 300	Gln	Lys	Thr	Glu
Ser 305	Gly	Ala	Gln	Ala	Leu 310	Gly	Asn	Glu	Leu	Val 315	Lys	Tyr	Leu	Lys	Ser 320
				325					330				Pro	335	
			340					345					Glu 350		
		3 55					360					365	Gln		
	370	_				375					380				Tyr
385				_	390					395					Glu 400
				405					410				Asp	415	
			420					425					Asp 430		
	_	435					440					445	Val		
	450					455					460		Val		
465					470					475					Asn 480
				485					490				Phe	495	
			500					505					Gln 510		
	_	51 5					520					525	Leu		
-	530					535					540		His		
545					550					555					Tyr 560
				565					570				Cys	575	
			580					585					Ala 590		
		59 5			_		600					605			Leu
	610	_				615					620				Glu
625	_				630					635					Glu 640
				645					650				Pro	655	
			660					665					Ala 670		
		675					680					685	Ser		
Lys	Asp	Leu	Leu	Gly	Ala	Ile	Ser	Gly	Ile	Ala	Ser	Pro	Tyr	Thr	Thr

	690					695					700				
Ser		Pro	Acn	Tle	Δen		Val	Arg	Δen	د ۱ ۵		Ser	Δτα	Gly	Ser
705	1111	110	ASH	110	710	JUL	V (4 1	71.9	ASII	715	тэр	JCI	711.9	Cry	720
	Tla	Car	Thr	λcn		Glv	λan	Ser	Len		Glu	λνα	λen	Car	
Leu	116	261	1111	725	Ser	GIY	ASII	261	730	PIO	GIU	Arg	ASII	735	Giu
T	C 0 22	7.00	Com		7.00	T	u: a	<i>(</i> 15)		C 0 25	002	TTID se	T 011		7 ~ ~
гуз	Ser	ASII		Leu	Asp	гуз	птз	Gln	GIII	ser	ser	IIII		GIY	ASII
_			740	~		. .	•	745	a 3 .	a .	~ 2	- 1	750		-
Ser	Val		Arg	Cys	Asp	гÀг		Asp	GIn	ser	GIU		ьуs	ser	Leu
_		755	_,	_	_		760	_	_		_	765	_		_
Leu		Cys	Pne	Leu	Tyr		Leu	Lys	Ser	Met		Asp	Asp	Ala	Leu
	770	_	_	_	_	775	_		_		780		_		
	Thr	Tyr	Trp	Asn	_	Ala	Ser	Thr	Ser		Leu	Met	Asp	Phe	
785	_			_	790					795			_		800
Thr	Ile	Ser	Glu		Cys	Leu	His	Gln		Gln	Tyr	Met	Gly	_	Arg
				805					810					815	
Tyr	Ile	Ala	Arg	Thr	Gly	Met	Met	His	Ala	Arg	Leu	Gln	Gln	Leu	Gly
			820					825					830		
Ser	Leu	Asp	Asn	Ser	Leu	Thr	Phe	Asn	His	Ser	Tyr	Gly	His	Ser	Asp
		835					840					845			
Ala	Asp	Val	Leu	His	Gln	Ser	Leu	Leu	Glu	Ala	Asn	Ile	Ala	Thr	Glu
	850					855					860				
Val	Cys	Leu	Thr	Ala	Leu	Asp	Thr	Leu	Ser	Leu	Phe	Thr	Leu	Ala	Phe
865					870					875					880
Lys	Asn	Gln	Leu	Leu	Ala	Asp	His	Gly	His	Asn	Pro	Leu	Met	Lys	Lys
-				885		_		_	890					895	-
Val	Phe	Asp	Val	Tyr	Leu	Cys	Phe	Leu	Gln	Lvs	His	Gln	Ser	Glu	Thr
		•	900	•		•		905		•			910		
Ala	Leu	Lvs		Val	Phe	Thr	Ala	Leu	Arg	Ser	Leu	Ile		Lvs	Phe
		915					920		-			925	•	1	-
Pro	Ser		Phe	Tvr	Glu	Glv		Ala	asa	Met	Cvs		Ala	Leu	Cvs
	930			-1-		935	5				940				-1-
Tvr		Ile	Leu	Lvs	Cvs		Asn	Ser	Lvs	Leu		Ser	Ile	Ara	Thr
945				-1-	950	-1-			-1-	955				5	960
	Δla	Ser	Gln	I.eu		Tvr	Phe	Leu	Met		Asn	Asn	Phe	Asn	
0_0			J	965		-1-			970	5				975	- / -
Thr	Glv	Lvs	Lve		Phe	Va 1	Δνα	Thr		Leu	Gln	Val	Tle		Ser
	017	Lys	980	561	1110	•••	**** 9	985			0111	vai	990		501
7 <i>7</i> = 7	Sar	Gln		Tla	Λla	Δen	₩ 1	Val	Glar	Tla	Gl v	Gly		λνα	Dhe
VAI	JCI	995	11-u	116	AIG	Asp	1000		Gry	110	GLY	1005		A. G	rne
Gln	Gln		T.011	Sar	Tla	Tla		Asn	Care	בומ	Λen			λνα	T 011
GIII	1010		Leu	261	116	1015		Maii	Cys	Ala	1020		Asp	Arg	Leu
Tlo			Thr	Cor	Dho			7	1751	T 1/6			Thr	T 1/0	A war
		nis	1111	Ser			ser	Asp	val		_	ьeu	1111	ьуѕ	_
1025		mh	37- 7	T 0	1030		mb	71-	~1 - -	1035		~1	TT-1	G1	1040
rre	Arg	Inr	val			Ата	inr	Ala			гуѕ	GIU	HIS		
_	_	~1		1045			.	~ 3	1050		_		_	1055	
Asp	Pro	GIU			val	Asp	ьeu	Gln	_	ser	ьeu	Ala	-		Tyr
_			1060					1065					1070		
Ala	Ser			Glu	Leu	Arg		Thr	Trp	Leu	Asp			Ala	Arg
		1075					1080					1085			
Ile	His	Val	Lys	Asn	Gly			Ser	Glu	Ala			Cys	Tyr	Val
	1090					1095					1100				
His	Val	Thr	Ala	Leu	Val	Ala	Glu	Tyr	Leu	Thr	Arg	Lys	Glu	Ala	Val
1105					1110					1115					1120
Gln	Trp	Glu	Pro	Pro	Leu	Leu	Pro	His	Ser	His	Ser	Ala	Cys	Leu	Arg

Arg Ser Arg Gly Gly Val Phe Arg Gln Gly Cys Thr Ala Phe Arg Val 1140					1125	=				1130	1				1139	5
1140	7	C	7 ~~~	~1			Dho	7. ~~	Cla			Thr	Δla	Phe		
The Thr	Arg	ser	Arg			Vai	Pne	ALG			Cys	1111	ΑΙα			Val
Met Gln		_,	_				<u>ما</u>	a 1			 .	11.4	~1			C1
Met Gln Asp Val His Phe Asn Glu Asp Val Leu Heu Tro Lus 1180 1180 1180 1180 1180 1180 1180 1180 1195 120 Asp Arg Tyr Glu Leu Ile Ala Ala Glu Tyr Glu Leu Ile Ala Glu Tyr Glu Leu Ala Asp Phe Ile Ile Tyr Glu Tyr Asp Asp Arg Arg <t< td=""><td>шe</td><td>Thr</td><td></td><td></td><td>iie</td><td>Asp</td><td>GIU</td><td></td><td></td><td>ser</td><td>Met</td><td>Mer</td><td></td><td></td><td>vai</td><td>Gry</td></t<>	шe	Thr			iie	Asp	GIU			ser	Met	Mer			vai	Gry
1170						_,	_				_				7	G1
Second Cys Ala Asp Gly Leu Trp Lys Ala Glu Arg Tyr Glu Leu Ile Ala 1185 1190 1195 1195 1195 1200 1205 1215 1210 1215	Met			Val	His	Phe			Asp	Val	Leu			Leu	Leu	GIU
1195																_ •
Asp Ile Tyr Lys Leu Ile Ile Pro Ile Tyr Glu Lys Arg Arg Asp Phe 1205	Gln	Cys	Ala	Asp	Gly	Leu	Trp	Lys	Ala	Glu			Glu	Leu	Ile	
1215																
Call Arg Leu Ala His Leu Tyr Asp Thr Leu His Arg Ala Tyr Ser Lys	Asp	Ile	Tyr	Lys	Leu	Ile	Ile	Pro	Ile	Tyr	Glu	Lys	Arg	Arg	Asp	Phe
Val																
Val	Glu	Arg	Leu	Ala	His	Leu	Tyr	Asp	Thr	Leu	His	Arg	Ala	Tyr	Ser	Lys
Arg																
Arg	Val	Thr	Glu	Val	Met	His	Ser	Gly	Arg	Arg	Leu	Leu	Gly	Thr	Tyr	Phe
1250																
1250	Arq	Val	Ala	Phe	Phe	Gly	Gln	Ala	Ala	Gln	Tyr	Gln	Phe	Thr	Asp	Ser
Simple S						-					-					
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Glu			Val	Glu	Glv	Phe	Phe	Glu	Asp	Glu	Asp	Gly	Lys	Glu	Tyr
The Tyr Lys Glu Pro Lys Leu Thr Pro Leu Ser Glu Ile Ser Gln Arg 1295 1295 1290 1295 1295 1295 1290 1295										_			•	•		
Leu Leu Lys Leu Tyr Ser Asp Lys Phe Gly Ser Glu Asp Ser Lys Met Lys Met Lys Lys Tyr Ser Gly Lys Ser Gly Ser Glu Asp Ser Lys Met Lys Tyr Lys Lys Tyr Lys Lys Tyr Lys Lys Lys Tyr Lys			Lvs	Glu	Pro			Thr	Pro	Leu	Ser	Glu	Ile	Ser	Gln	Arq
Leu Leu Lys Leu Tyr Ser Asp Lys Phe Gly Ser Glu Asn Val Lys Met		- / -	2,0													
130 130	T.011	Len	T.VC	Leu			Δsp	Lvs	Phe			Glu	Asn	Val	Lvs	Met
The Gln	neu	Lea	275			501		-,-								
Mathematical Registration Mathematical Registration	т1 о	Cln	7.00			Lvc	v-1	Aen			Δen	Len	Asn			Tvr
Ala Tyr Ile Gln Val Thr His Val Ile Pro Phe Asp Glu Lys Glu 1330	116	GIII	_		OLY	Lys	Val			2,5	1105				-1-	-1-
Leu Gln Glu Arg Lys Thr Glu Phe Glu Arg Ser His Asn Ile Arg Arg 1345	ה I ה	The race			17 = 1	Thr	Uic			Pro	Dhe	Dhe			Lvs	Glu
Leu Gln Glu Arg Lys Lys Thr Glu Phe Glu Arg Ser His Asn Ile Arg Arg Asn Ile Arg Arg Arg 1365 1355 1360 Phe Glu Arg Ile Arg Arg Gln Gly Gly Gly Gly Gly Ile Arg Arg Gln Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly	AIA	-		GIII	Val	1111			110	110	1 110				-1-	
1345	T 011			λνα	Lve	Thr			Glu	Δrα	Ser			Tle	Ara	Ara
Phe Met Phe Glu Met Pro Phe Thr Gln Thr Gly Lys Arg Gln Gly Gly 1365			JIU	nrg	כעם					•••=					5	
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$			Dhe	Glu	Mat			Thr	Gln	Thr			Δrα	Gln	Glv	
Val Glu Glu Glu Gln Cys Lys Arg Arg Arg Thr Ile Leu Thr Ala Ile His Cys 1380 1385 1390 Phe Pro Tyr Val Lys Lys Lys Arg Ile Pro Val Met Tyr Gln His His Thr 1395 1400 1405 Asp Leu Asn Pro Ile Glu Val Ala Ile Asp Glu Met Ser Lys Lys Val 1410 1415 1420 Ala Glu Leu Arg Gln Leu Cys Ser Ser Ala Glu Val Asp Met Ile Lys 1425 1430 1435 1440 Leu Gln Leu Lys Leu Gln Gly Ser Val Ser Val Gln Val Asn Ala Gly 1450 Pro Leu Ala Tyr Ala Arg Ala Phe Leu Asp Asp Thr Asn Thr Lys Arg 1460 Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe 1475 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Ile Lys	FIIC	Mec	FILE	Oru			1110		0111			-,-				
1380	37-1	Glu	Glu	Gln			Δνα	Ara	Thr			Thr	Δla	Tle		
Phe Pro Tyr Val Lys Lys Arg Ile Pro Val Met Tyr Gln His His Thr 1395	vai	GIU	GIU			БуЗ	9	ni 9			DC u	1				-7-
Asp Leu Asn Pro Ile Glu Val Ala Ile Asp Glu Met Ser Lys Lys Val 1415 Ala Glu Leu Arg Gln Leu Cys Ser Ser Ala Glu Val Asp Met Ile Lys 1425 Leu Gln Leu Lys Leu Gln Gly Ser Val Gln Val Asp Met Ile Lys Leu Gln Leu Cys Ser Val Gln Val Asp Met Ile Lys 1435 Fro Leu Ala Tyr Ala Arg Ala Phe Leu Asp Asp Thr Asn Thr Lys Arg 1455 Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe 1475 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Lys Clu Val Phe Arg Cln Phe Lys Val Glu Ala Cys Cys Cys Cys Cys Cys Cys Cys Cys Cys	Dho	Dre	Tire			Tvc	λνα	Tla			Mot	Туг	Gln			Thr
Asp Leu Asn Pro Ile Glu Val Ala Ile Asp Glu Met Ser Lys Lys Val 1410 Ala Glu Leu Arg Gln Leu Cys Ser Ser Ala Glu Val Asp Met Ile Lys 1420 Leu Gln Leu Lys Leu Gln Gly Ser Val Gln Val Asp Met Ile Lys 1440 Leu Gln Leu Lys Leu Gln Gly Ser Val Gln Val Asp Ala Gly 145 Pro Leu Ala Tyr Ala Arg Ala Phe Leu Asp Asp Thr Asn Thr Lys Arg 1460 Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe 1475 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Ile Lys	PHE	PLO	-		пуъ	Буз	nrg			Val	Mee	- y -				
Ala Glu Leu Arg Gln Leu Cys Ser Ser Ala Glu Val Asp Met Ile Lys 1425	* ~~	T			Tla	C1.1	Val			λen	G111	Met			Lvs	Val
Ala Glu Leu Arg Gln Leu Cys Ser Ser Ala Glu Val Asp Met Ile Lys 1425	Asp			PLO	116	Gru			116	тэр	O1u			ביים	,	,
1425				3	<u>ما</u>	T			Cor	77-	C1.,			Mat	Tla	Lve
Leu Gln Leu Lys Leu Gln Gly Ser Val Ser Val Gln Val Asn Ala Gly 1445 1450 1455 Pro Leu Ala Tyr Ala Arg Ala Phe Leu Asp Asp Thr Asn Thr Lys Arg 1460 1465 1470 Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe 1475 1480 1485 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Ile Lys			Leu	Arg	GIII			ser	261	Ата			rab	HEC	110	
Pro Leu Ala Tyr Ala Arg Ala Phe Leu Asp Asp Thr Asn Thr Lys Arg 1465 Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe 1475 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Lys Clu Lys Clu Arg Clu Lys Clu Arg Clu Lys Clu Arg Clu Arg Clu Lys Clu Arg Clu Arg Clu Lys Clu Arg Clu Lys Clu Arg Clu Arg Clu Lys Clu Arg Clu Arg Clu Lys Clu Arg Clu Arg Clu Arg Clu Lys Clu Arg Clu Ar			*	T	T			C	17-7	C 0 20			1701	Acn	ת 1 ת	
Pro Leu Ala Tyr Ala Arg Ala Phe Leu Asp Asp Thr Asn Thr Lys Arg 1460 1465 1470 Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe 1475 1480 1485 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Ile Lys	Leu	GIn	Leu	ьуѕ			GIY	ser	val			GIII	Val	ASII		
Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe 1475 1480 1485 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Lys	_	_					3.1 -	D1	.			mla sa	7	mb se		
Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe 1475 1480 1485 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Ile Lys	Pro	Leu	Ala	_		Arg	Ala	Pne			Asp	Thr	ASI			Arg
1475 1480 1485 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Ile Lys							_	_					-1			D1
Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Ile Lys	Tyr	Pro			Lys	Val	Lys			Lys	GIu	Val			GTU	рпе
										_						_
1490 1495 1500	Val	Glu	Ala	Cys	Gly	Gln			Ala	Val	Asn			Leu	Ile	Lys
								_								
Glu Asp Gln Leu Glu Tyr Gln Glu Glu Met Lys Ala Asn Tyr Arg Glu	Glu	Asp	Gln	Leu	Glu			Glu	Glu	Met			Asn	Tyr	Arg	
1505 1510 1515 1520																
Met Ala Lys Glu Leu Ser Glu Ile Met His Glu Gln Ile Cys Pro Leu	Met	Ala	Lys	Glu	Leu	Ser	Glu	Ile	Met	His	Glu	Gln	Ile	Cys	Pro	Leu
1525 1530 1535																
Glu Glu Lys Thr Ser Val Leu Pro Asn Ser Leu His Ile Phe Asn Ala	Glu	Glu	Lys	Thr	Ser	Val	Leu	Pro	Asn	Ser	Leu	His	Ile	Phe	Asn	Ala
1540 1545 1550				1540)				1545	5				1550)	
Ile Ser Gly Thr Pro Thr Ser Thr Met Val His Gly Met Thr Ser Ser		C ~ ~	Glv	Thr	Pro	Thr	Ser	Thr	Met	Val	His	Gly	Met	Thr	Ser	Ser

```
1560
                                                1565
        1555
Ser Ser Val Val
   1570
<210> 159
<211> 540
<212> DNA
<213> Homo sapiens
<400> 159
geoggetetg ceatgtgett actotgagee acctaacete ggegtgette agtttactea
tecqctcate tqcagaatgg gtgatgctgt eggtactteg tggcatacag gaaagtgeee
agcatggtca gcctcagtga gaggtggcca gtggggagtg gtggccactg tacacctggc
180
acaqcccaga gatgcatgtg ccactetgtt gtgtgcttca accaaggggc gctctggcag
ggcttgggtg ggacttccca aagggcatgg aaaagttccc agtcaatgag atccatggag
300
acceatggga gtgggggtea geeccageet aagaggaeee ceageeetge eetgtgeeee
aggacacace aggeactgte cettgtegee tteccagaca acctgtacce tecaggeeae
cagttctcgt ccatgacaaa gaaaggagcc ttctaaataa gtgcccgcca gaggctgcac
getteeetge eeetteeggg tggaeetggg ttteaaagag aagetgeeag tgeaacgegt
<210> 160
<211> 110
<212> PRT
<213> Homo sapiens
<400> 160
Met Val Ser Leu Ser Glu Arg Trp Pro Val Gly Ser Gly Gly His Cys
                                    10
1
Thr Pro Gly Thr Ala Gln Arg Cys Met Cys His Ser Val Val Cys Phe
Asn Gln Gly Ala Leu Trp Gln Gly Leu Gly Gly Thr Ser Gln Arg Ala
                            40
Trp Lys Ser Ser Gln Ser Met Arg Ser Met Glu Thr His Gly Ser Gly
                        55
Gly Gln Pro Gln Pro Lys Arg Thr Pro Ser Pro Ala Leu Cys Pro Arg
                                        75
                    70
Thr His Gln Ala Leu Ser Leu Val Ala Phe Pro Asp Asn Leu Tyr Pro
                85
                                    90
Pro Gly His Gln Phe Ser Ser Met Thr Lys Lys Gly Ala Phe
                                                    110
            100
                                105
<210> 161
<211> 351
<212> DNA
<213> Homo sapiens
```

```
<400> 161
nnacgcgtac gtctttcggc cgaagaagga acgtgggcag gggcctcctt cgctggccgc
cgcgcttggc tcgcagcgac gatgaagggc gacgacagca gcaagatcac ccacaagatc
120
qcccqqqcqa agcgcgaggg ccgcgtatgg tggagctttg agtacttccc gccgcgcacg
ccqcaqqqca tqcaqaattt gtatqaccqt atcqaqcqca tqaqtcaqct gggccccqaq
tttgtggaca ttacgtggaa tgccgggggc cggacgtcgg atatgacgac gcagctggtc
aaqacqqtqc atgcgtactt tggtgtcgag acgtgcatgc atctgacgtg c
351
<210> 162
<211> 117
<212> PRT
<213> Homo sapiens
<400> 162
Xaa Arg Val Arg Leu Ser Ala Glu Glu Gly Thr Trp Ala Gly Ala Ser
                                    10
Phe Ala Gly Arg Arg Ala Trp Leu Ala Ala Thr Met Lys Gly Asp Asp
                                25
Ser Ser Lys Ile Thr His Lys Ile Ala Arg Ala Lys Arg Glu Gly Arg
                            40
Val Trp Trp Ser Phe Glu Tyr Phe Pro Pro Arg Thr Pro Gln Gly Met
                        55
Gln Asn Leu Tyr Asp Arg Ile Glu Arg Met Ser Gln Leu Gly Pro Glu
                                        75
                    70
Phe Val Asp Ile Thr Trp Asn Ala Gly Gly Arg Thr Ser Asp Met Thr
                                    90
Thr Gln Leu Val Lys Thr Val His Ala Tyr Phe Gly Val Glu Thr Cys
                                105
            100
Met His Leu Thr Cys
        115
<210> 163
<211> 360
<212> DNA
<213> Homo sapiens
<400> 163
gegtgeteca teggeacett geagatggge gaattegetg aaaaegtege eggtggegte
gacacctaca ecetgegtea geceategge gtatgegeag geateactee gtteaactte
coggogatga ttocactgtg gatgttoccg atggcgattg cotgcggtaa cactttogtg
ctcaaaccgt ccgaacaaga ccctctgtcg acgatgctgc tggtagaact ggcgctggaa
geeggtgtge eggeeggegt geteaaegtg gtgeaeggeg geaaggatgt ggtggatgeg
300
```

```
ctgtgcaccc ataaagatat caaggcagtt tctttcgtcg gttcgaccgc cgttggtacc
360
<210> 164
<211> 120
<212> PRT
<213> Homo sapiens
<400> 164
Ala Cys Ser Ile Gly Thr Leu Gln Met Gly Glu Phe Ala Glu Asn Val
                                    10
Ala Gly Gly Val Asp Thr Tyr Thr Leu Arg Gln Pro Ile Gly Val Cys
                                25
            2.0
Ala Gly Ile Thr Pro Phe Asn Phe Pro Ala Met Ile Pro Leu Trp Met
                                                45
                            40
Phe Pro Met Ala Ile Ala Cys Gly Asn Thr Phe Val Leu Lys Pro Ser
                                            60
                        55
    50
Glu Gln Asp Pro Leu Ser Thr Met Leu Leu Val Glu Leu Ala Leu Glu
                                        75
                    70
Ala Gly Val Pro Ala Gly Val Leu Asn Val Val His Gly Gly Lys Asp
                                    90
Val Val Asp Ala Leu Cys Thr His Lys Asp Ile Lys Ala Val Ser Phe
                                105
            100
Val Gly Ser Thr Ala Val Gly Thr
                            120
        115
<210> 165
<211> 728
<212> DNA
<213> Homo sapiens
<400> 165
gctagcagcc ttcaccctcc tagaggggca ggctcggcga caaggggcgg gggtgccccg
teccagegag ggacgecegg ggetgggggt geeggtegag eeeggggeaa eagetteace
aagtttggca accgcaacgt cttcatgaag gacaacagct cttcttccag cacagactcc
egeteceget cetectecag gteceegacg egecaettee geagaagtga eteceaetea
240
gactocgaca getectacte agggaatgag tgtcaccetg tgggeegeag gaaccegeee
cctaagggcc ggggcggtcg aggggcccat atggatcggg gccgaggcag ggcgcagcgt
gggaagaggc acgatctggc gcccaccaag cgcagtcgaa agaagatggc ggcgctggag
tgtgaggacc cggagcgaga gctgaagaag cagaagcggg cagcccgctt ccagcacgga
cactecegee geetgegeet egageeeetg gtgetgeaga tgageageet ggagageagt
ggggctgacc ctgactggca ggagctgcag atcgtgggca cctgccctga catcaccaag
cactacetge geeteacetg tgeeceegae eegteeaceg tgegeeetgt ggeatteeet
660
```

```
gtggcaggtt ttgaaaaagt cgctgtgcat ggtcaagtgc cactggaaag agaagcagga
720
ctacgcgt
728
<210> 166
<211> 242
<212> PRT
<213> Homo sapiens
<400> 166
Ala Ser Ser Leu His Pro Pro Arg Gly Ala Gly Ser Ala Thr Arg Gly
                                   10
Gly Gly Ala Pro Ser Gln Arg Gly Thr Pro Gly Ala Gly Gly Ala Gly
                               25
Arg Ala Arg Gly Asn Ser Phe Thr Lys Phe Gly Asn Arg Asn Val Phe
                           40
Met Lys Asp Asn Ser Ser Ser Ser Ser Thr Asp Ser Arg Ser Arg Ser
                                           60
                       55
Ser Ser Arg Ser Pro Thr Arg His Phe Arg Arg Ser Asp Ser His Ser
                   70
Asp Ser Asp Ser Ser Tyr Ser Gly Asn Glu Cys His Pro Val Gly Arg
                                   90
               85
Arg Asn Pro Pro Lys Gly Arg Gly Arg Gly Ala His Met Asp
                               105
Arg Gly Arg Gly Arg Ala Gln Arg Gly Lys Arg His Asp Leu Ala Pro
                           120
Thr Lys Arg Ser Arg Lys Lys Met Ala Ala Leu Glu Cys Glu Asp Pro
                       135
                                           140
Glu Arg Glu Leu Lys Lys Gln Lys Arg Ala Ala Arg Phe Gln His Gly
                   150
                                       155
His Ser Arg Arg Leu Arg Leu Glu Pro Leu Val Leu Gln Met Ser Ser
               165
                                  170
Leu Glu Ser Ser Gly Ala Asp Pro Asp Trp Gln Glu Leu Gln Ile Val
                              185
Gly Thr Cys Pro Asp Ile Thr Lys His Tyr Leu Arg Leu Thr Cys Ala
                           200
Pro Asp Pro Ser Thr Val Arg Pro Val Ala Phe Pro Val Ala Gly Phe
                       215
Glu Lys Val Ala Val His Gly Gln Val Pro Leu Glu Arg Glu Ala Gly
225
                   230
Leu Arg
<210> 167
<211> 510
<212> DNA
<213> Homo sapiens
<400> 167
nnacqcqtqq aaccagaact caggcccqtg tgaggagtct ggtttggaac acacggggcc
qcaacacaga attgtcaggt cctgtgccgt gaccaccacc cctcgggcca tgccaggtgc
120
```

```
tggtgagggg caggtggctc ccgccaggcg cctgctggcc tgaccgcact ccgtccacag
180
gtecteatgg gegtecteeg getgggette gtgteegeet aceteteaca gecaetgete
240
gatggctttg ccatgggggc ctccgtgacc atcctgacct cgcagctcaa acacctgctg
ggegtgegga teeegeggea eeaggggeee ggeatggtgg teeteacatg getgageetg
360
ctgcgcggcg ccgggcaggc caacgtgtgc gacgtggtca ccagcacggt gtgcctggcg
gtgctgctag ccgcgaagga gctctcagac cgctaccgac accgcctgag ggtgccgctg
cccacggagc tgctggtcat cgtggtggcc
510
<210> 168
<211> 128
<212> PRT
<213> Homo sapiens
<400> 168
Gly Ala Gly Gly Ser Arg Gln Ala Pro Ala Gly Leu Thr Ala Leu Arg
Pro Gln Val Leu Met Gly Val Leu Arg Leu Gly Phe Val Ser Ala Tyr
                                25
            20
Leu Ser Gln Pro Leu Leu Asp Gly Phe Ala Met Gly Ala Ser Val Thr
                            40
Ile Leu Thr Ser Gln Leu Lys His Leu Leu Gly Val Arg Ile Pro Arg
                                            60
His Gln Gly Pro Gly Met Val Val Leu Thr Trp Leu Ser Leu Leu Arg
                                        75
                    70
Gly Ala Gly Gln Ala Asn Val Cys Asp Val Val Thr Ser Thr Val Cys
                                    90
Leu Ala Val Leu Leu Ala Ala Lys Glu Leu Ser Asp Arg Tyr Arg His
                                105
Arg Leu Arg Val Pro Leu Pro Thr Glu Leu Leu Val Ile Val Val Ala
                                                 125
                            120
        115
<210> 169
<211> 537
<212> DNA
<213> Homo sapiens
<400> 169
gaattccacc gcatgtcgtg tctggacgta tgtaggtcgc ggtagtgtgc gaccgccggt
gccttaaagg agagcgggca tcggcgttgc agtacgagag gggaaggtgt gcggatactt
attgtcggtg cggcatcgtc catccacacc gttcgatggg tcaatggact ggtcaagcgg
ggtcacgagg ttcacctggc atcagtccat ccggcgggcc gtcactccat tgatccccga
gttcggatcc acctggcccc acacggcggg aaggcaaaat acgtcgtcaa tgccggctgg
300
```

```
ctgcgatcag tggcggctgg ggtgcaacct gacatcgtca acgtccacta tgcgaccggt
tatggtctgc tcgctcgtct tgcccatatt gacgccccga cgctgctgtc ggtgtgggga
aqtgacgttt acgattcccc ccgggcaaat cccctcatgc gtcacatggt ccgatccaac
ttggtctcag ctactcggat cgcatcgaca agccactgca tggcgcgtgt cacgcgt
537
<210> 170
<211> 164
<212> PRT
<213> Homo sapiens
<400> 170
Cys Ala Thr Ala Gly Ala Leu Lys Glu Ser Gly His Arg Arg Cys Ser
Thr Arg Gly Glu Gly Val Arg Ile Leu Ile Val Gly Ala Ala Ser Ser
            20
                                25
Ile His Thr Val Arg Trp Val Asn Gly Leu Val Lys Arg Gly His Glu
                            40
                                                45
Val His Leu Ala Ser Val His Pro Ala Gly Arg His Ser Ile Asp Pro
                        55
Arg Val Arg Ile His Leu Ala Pro His Gly Gly Lys Ala Lys Tyr Val
                                        75
                    70
Val Asn Ala Gly Trp Leu Arg Ser Val Ala Ala Gly Val Gln Pro Asp
                85
Ile Val Asn Val His Tyr Ala Thr Gly Tyr Gly Leu Leu Ala Arg Leu
                                105
Ala His Ile Asp Ala Pro Thr Leu Leu Ser Val Trp Gly Ser Asp Val
                                                125
                            120
Tyr Asp Ser Pro Arg Ala Asn Pro Leu Met Arg His Met Val Arg Ser
                                            140
                        135
Asn Leu Val Ser Ala Thr Arg Ile Ala Ser Thr Ser His Cys Met Ala
                    150
                                        155
145
Arg Val Thr Arg
<210> 171
<211> 391
<212> DNA
<213> Homo sapiens
<400> 171
ctagacaagc tegegegggt gggettegae actettggte tacagacett ectaactgeg
ggggagaagg agtcccgcgc atggacgatt cacaagggcg acaccgcccc tgaggctgct
ggcgtcatcc ataccgactt ccagaagggg ttcatcaagg cccaggtggt gtccttcggc
180
gacettgttg aatttggegg egaaaaggag geccaggetg etgggaaget geggttggag
ggcaaggagt acgttatgca ggacggtgac gtagtggaat tccgatttaa cgtgtagctc
300
```

```
tggtttgata cttacttggc ttaaccgcat ctgagatccg tcatatcttt ggcgtagcct
tattggtatg aataacatgc cgtagccaaa g
391
<210> 172
<211> 98
<212> PRT
<213> Homo sapiens
<400> 172
Leu Asp Lys Leu Ala Arg Val Gly Phe Asp Thr Leu Gly Leu Gln Thr
Phe Leu Thr Ala Gly Glu Lys Glu Ser Arg Ala Trp Thr Ile His Lys
                                25
Gly Asp Thr Ala Pro Glu Ala Ala Gly Val Ile His Thr Asp Phe Gln
        35
                            40
Lys Gly Phe Ile Lys Ala Gln Val Val Ser Phe Gly Asp Leu Val Glu
                        55
Phe Gly Gly Glu Lys Glu Ala Gln Ala Gly Lys Leu Arg Leu Glu
                                        75
                    70
Gly Lys Glu Tyr Val Met Gln Asp Gly Asp Val Val Glu Phe Arg Phe
                85
                                    90
Asn Val
<210> 173
<211> 309
<212> DNA
<213> Homo sapiens
<400> 173
ccatggagtg tecettgtgc gagcattttg agagctatac caacacccat ccctgcaggt
cccagagccg agccatttct caggagagca ggaagggagc aggccgaggg gtgctcccag
ccageccegg aaccegaggt ctggggacge agcegaccag ccetecttgt ctgggcetet
gtttcctctt cgacacaggg aagcagggag gggccgatca gcgacttagg cctgttggct
gtggtggggt cccctgcgtt tctgggaagc cacggaccct gggatgtacc tgggtttcat
300
tcgcagtga
309
<210> 174
<211> 102
<212> PRT
<213> Homo sapiens
<400> 174
Met Glu Cys Pro Leu Cys Glu His Phe Glu Ser Tyr Thr Asn Thr His
                                    10
Pro Cys Arg Ser Gln Ser Arg Ala Ile Ser Gln Glu Ser Arg Lys Gly
```

30

25 20 Ala Gly Arq Gly Val Leu Pro Ala Ser Pro Gly Thr Arg Gly Leu Gly 40 Thr Gln Pro Thr Ser Pro Pro Cys Leu Gly Leu Cys Phe Leu Phe Asp Thr Gly Lys Gln Gly Gly Ala Asp Gln Arg Leu Arg Pro Val Gly Cys 70 65 75 Gly Gly Val Pro Cys Val Ser Gly Lys Pro Arg Thr Leu Gly Cys Thr 85 90 Trp Val Ser Phe Ala Val 100 <210> 175 <211> 8484 <212> DNA <213> Homo sapiens <400> 175 nnactttttt tttttttt catttatgct atggagaaac cagcatggag atgtcatggg agageatgea caggeeeege eetagggagt ggtgatgtgt ttggggaggt gettgtttee aggtecated cacacgttgt coagttggat cotatggcag gotggctgtg gotttetete 180 tectgettet ettecteete cagataaggg tetgeaggat ettetgetta geaagtggtg gccaaggact ggtggatggg tggctggaag cagcgcacat gctccacagt ggaactgtct gtctccacgg acttcatgta tttgttcagg atggcaaaaa cctcattgtt caagatctga 360 tacttcctga tccggtcggc catcttcttc aggggcacat tcttaatgat ttcatccttc ccgtcctgcc tctgcacttt tagcaggtgg taacagaagt cgaacaggtc aaagcgacgc tgetggeeca geaggacaat gatggagcaa ceageceagt teaagecate geegaaacae tgctcagctg tgaactcgtt ggttcccaca gggatgcagt acacgaactg catggcgctc 600 cacagooggt ggaactocac acacteatog acgtgcatga cgccattggt gggcggtggg ccccgccaga tggggtcctg caggtagctc cgaatgcggg tcaggatgac ctcaaacatg 720 gacaggecae ageacageeg eteettggte aggaggteae eetegegage aatggegatt tgctgagggg tccccagccg ctcgatcaga gggaccaggt ggagcggggc atacttggct tecagacgtt teattttgge ateaagtete tececetett teacatggae tegeggeaag atgttctgga aaggagccgc gtgcagcagg tcacacactt cttctaaaga cagagctctg ctcgatgagg aggcagaaga ggatggcatt gcccacttcc ctcaggctct ggaagcacgt ctgttttgag ctctgcgtac tcaatgatgt ccttcagctg gtggtggaag aactccagga 1080

tccctgggga 1140	gccatactca	tgtcggggca	agcggcatat	cttgggcatc	acctctatca
	atactggaga	atggttcctt	ggagcaagct	cttcacaatc	tttagcagtt
cctccatgac 1260	cacagegatg	ccctgataac	ccaggagtct	gcagatagtc	ttgaaagtgt
ggaggtccca 1320	cgaagttccg	gtagctgccg	taaatgctgg	agtaggccaa	gttcaaagcc
1380	gcagatactg				
1440	acactgtccg				
1500	agttgagctc				
1560	tggcctcccg				
1620	gggtcatgcg				
1680	tttcaaatcg				
1740	tgatcagacg				
1800	gcgtctcgta				
1860	atcgtaaccg				
1920	ggtctgccag				
1980	cgtacaggaa				
2040	ccagggagta				
2100	ttagaatcca				
2160	ccaggaagaa				
2220	tgaaatttat				
2280	caatgggtcc				
2340	gtgattcaag				
2400	ggggcacctt				
2460	tagggggctc				
2520	tgatgaggac				
2580	gggcgaagtc				
2640	ccatcctgcc				
acgaaggcaa 2700	atttttcctc	actggtgtaa	regragegig	rggetetete	acallected

gcggtgccag 2760	gacagtcctt	gttgcagaac	ttgtctgtgg	gatgaaccag	cttccaagag
tacacctcca 2820	tgacgtgggc	gctccacttg	gatagaagct	gcagaccccg	cagggctagg
tcgaagagct 2880	cgcgatactc	ctcgtctgac	ttctggctgt	ccagccctga	gcccgtcacc
acctcactgt 2940	tgctgtagcg	agcgagctcg	gagatgaagc	ggatgtggtc	ateceggate
tgaaccatct 3000	gctcgcagat	attgtactgg	gggctgatgc	tgctctgggt	gcacgtccat
cgagatttat 3060	tttcctcgta	gtgggcgctg	gtcttgatat	atcttgccag	ttctatttgc
atgtccccaa 3120	atagtggaac	cacctggagt	tgcttgaagt	acttgtcgat	tttggataag
tttattcttt 3180	tcttggcatc	caacttatag	atgttactga	cactcccatc	catcaggtac
agaccaaatc 3240	ccatgacttt	gagaagcatg	tgtttctcac	tgggcgtcaa	atacatcctg
ttctcgtagt 3300	aatccacaca	cagattcaca	atatctgcca	ggagctcttc	gtagccagaa
3360				tatgattggc	
3420				ttttacgtaa	
gccctcttgt 3480	aggcagagtg	gtcgttcttg	acgetgeact	tcatgttctt	cagategtee
3540				ggtaggcttc	
3600				tgcagaaccg	
3660				ccggctccag	
3720				atttcacctg	
3780				catgteette	
3840				ttgcaatgcc	
3900				acatgatgga	
3960				caagcaggtc	
4020				tttctgtatt	
4080				ggccccgcgc	
4140				cagaaactgc	
4200				tgcttgaaga	
4260				tcatgtacca	
gacacaaact 4320	ttgaggacag	gaatgcattt	gtcacgggca	ttgcaaggta	cattgagcag

gctacagtcc 4380	actccagcat	gaatgagatg	ctggaggaag	gacatgagta	tgcggtcatg
ctgtacacct	ggcgcagctg	ctcccgggcc	atcccacagg	tgaaatgtaa	cgagcagcct
aacagagtgg 4500	aaatctacga	gaaaaccgtg	gaggttctgg	agcctgaggt	cacaaaactg
	tgtacttcca	gagaaatgcc	attgagcgtt	tctgcgggga	agtgaggcgc
	ccgagaggag	gaaggacttc	gtgtcagaag	cctacctgat	cacactgggc
	acatgttcgc	tgtgctggac	gagctgaaga	acatgaagtg	cagtgtgaag
aacgaccact 4740	cagcgtacaa	gagggccgct	cagtttttac	gtaaaatggc	agatccacag
tccatccagg 4800	aatcgcagaa	tctgtccatg	ttcctggcca	atcataacaa	gatcacacag
tctctgcagc 4860	agcagctcga	agtgatttct	ggctacgaag	agctcctggc	agatattgtg
aatctgtgtg 4920	tggattacta	cgagaacagg	atgtatttga	cgcccagtga	gaaacacatg
cttctcaaag 4980	tcatgggatt	tggtctgtac	ctgatggatg	ggagtgtcag	taacatctat
aagttggatg 5040	ccaagaaaag	aataaactta	tccaaaatcg	acaagtactt	caagcaactc
caggtggttc 5100	cgctatttgg	ggacatgcaa	atagaactgg	caagatatat	caagaccagc
gcccactacg 5160	aggaaaataa	atctcgatgg	acgtgcacat	cctccggcag	cagccctcag
tacaacatct 5220	gcgagcagat	gatccagatc	cgcgaggacc	acatgcgctt	catttcggag
ctggcgcgct 5280	acagcaacag	cgaggtggtc	acgggctcgg	gccgccagga	ggcccagaag
acggacgcgg 5340	agtaccgcaa	gctcttcgac	ctggcgctgc	agggcctgca	gctgttgtcg
cagtggagcg 5400	cgcacgtgat	ggaagtgtat	tcctggaagc	ttgtgcaccc	caccgacaag
tactccaaca 5460	aggactgccc	cgacagcgct	gaagagtacg	agegtgecae	gcgctacaac
tacaccagcg 5520	aggagaagtt	tgccctagtg	gaggtgatcg	ccatgatcaa	aggeetgeag
5580		gagcgtgttc			
gcactgcagg 5640	acttctccca	ggtgaccctt	agggagccgc	tgeggeagge	catcaagaag
aagaagaacg 5700	tcatccagag	tgtcctgcag	gccatcagga	agaccgtgtg	tgactgggag
acggggcatg 5760	agcccttcaa	tgacccagcc	ttgcggggcg	agaaggaccc	caagagegge
ttcgacataa 5820	aagtaccacg	ccgcgccgtg	ggaccctcca	gcactcagct	ttacatggtg
5880		cattgcagac			
agccttgagg 5940	ggcccaccat	attggacata	gaaaaattţc	atcgagagtc	attcttctac

actcacttga 6000	taaatttcag	tgaaacgctg	cagcagtgct	gtgacctttc	gcagctgtgg
ttccgagagt 6060	tcttcctgga	gctgaccatg	ggcaggagga	tccagttccc	cattgagatg
tcgatgccct 6120	ggatcctgac	ggaccacatc	ctggagacca	aggaggcatc	gatgatggag
6180			gacagcgccc		
6240			gccgaggtga		
6300			tattataagg		
6360			aatcagggag		
6420			aggcatgtgc		
6480			tcagcagcca		
6540			acctccatag		
6600			agccggtacc		
6660			tcagcgccct		
6720	_		cccaactact		
6780			caggaatttc		
6840			gctttgaact		
6900			cactttcaag		
6 96 0			ctgctgaagg		
7020			atggaggtga		
7080			ctggagttct		
7140			tgcttccaga		
7200			agcctgtctt		
7260			ccgcgagtcc		
7320			aagtacgccc		
7380			gccatcgcaa		
7440			tttgaggtca		
7500			ctgcccagca		
gagtgtgtgg 7560	agtttcacag	actgtggagt	gccatgcagt	ttgtctactg	cattcccgtg

gggacacacg agttcacagt cgagcagtgc tttggtgatg ggctacactg ggctggctgt atgatcatcg tacttcttgg gcagcagcgg cgttttgctg tgctggattt ctgctaccat 7680 ctacttaaag tccagaaaca tgatggcaaa gatgagatta ttaaaaaatgt gcctttgaag aagatggtgg agagaattcg caagttccag attctcaatg atgagatcat caccatcctg gataagtacc tgaagtcagg cgacggggag ggcacgccag tggagcatgt gcgctgcttc cagoogocca tocaccagte cotogocage agotgaggge acgogotgca otoogtaact caacatggca tgcctttctc tccgtaaact atttagtgag atttttaggg actatttttc agtatetetg tacetgttaa agggggtget tttegateta aaaaettaat tttataaaaat tgacttattt ttctagacta aaattgtata tgcttttggt aattaggaac tcttgagaat attggctgct gattgttgcc atcacgttcc tacaaaattg tttttctatg ggatgttctg 8160 gcagctgtgt cataaaatgc tgctgggttc attcattcat tccataagaa acttaatacc aqcaaatgca ttaaatccct tgccagttac cattaactgt aactatttag cttttgttta gggatettte tgatggtett ttatgageaa tettagttet aagteattgt teecateeet 8340 tttttgtgtg tttcagaaaa tagtgaactt gattcccctg cttccactaa atccagttgt gacaaaatct aacgtgacat cagatcgaaa ggttatagaa ataaaactaa tgagatctaa aaaaaaaaa aaaaaaaaaa aaaa 8484 <210> 176 <211> 1393 <212> PRT <213> Homo sapiens <400> 176 Met Ala Arg Glu Gln Leu Arg Gln Val Tyr Ser Met Thr Ala Tyr Ser 10 15 Cys Pro Ser Ser Ser Ile Ser Phe Met Leu Glu Trp Thr Val Ala Cys 20 25 Ser Met Tyr Leu Ala Met Pro Val Thr Asn Ala Phe Leu Ser Ser Lys 40 Phe Val Ser Lys Leu Ala Trp Tyr Met Met Glu Glu Gly Gly Gly Ser 55 60 Met His Gly Cys Trp Ser Gly Arg Gly Ser Ser Ser Ser Arg Ser Thr 70 75 Leu Asp Arq Ala Ser Ser Arq Val Thr Cys Val Val Met Ala Ala Val 85 90 Ser Val Phe Cys Thr Gly Ser Ala Ala Gly Pro Gly Glu Gly Pro Glu 105 Ala Thr Ala Gly Pro Arg Ala Gly Ala Gln Asp Ala Leu Pro Arg Ser

		115					120					125			
Ala	Ala		Pro	Val	Gln	Asn			Thr	Ala	Ala		Thr	Thr	His
	130					135					140				
Val		Leu	Glu	Asp	Ala		Ser	Asn	Val	Asp		Leu	Glu	Glu	Leu
145				-	150					155					160
Pro	Leu	Pro	Asp	Gln	Gln	Pro	Cys	Ile	Glu		Pro	Pro	Ser	Ser	Ile
			-	165			•		170					175	
Met	Tyr	Gln	Ala	Asn	Phe	Asp	Thr	Asn	Phe	Glu	Asp	Arg	Asn	Ala	Phe
			180					185					190		
Val	Thr	Gly	Ile	Ala	Arg	Tyr	Ile	Glu	Gln	Ala	Thr	Val	His	Ser	Ser
		195					200					205			
Met	Asn	Glu	Met	Leu	Glu	Glu	Gly	His	Glu	Tyr	Ala	Val	Met	Leu	Tyr
	210					215					220				
	Trp	Arg	Ser	Cys		Arg	Ala	Ile	Pro		Val	Lys	Cys	Asn	
225					230			_		235					240
Gln	Pro	Asn	Arg	Val	Glu	Ile	Tyr	Glu	-	Thr	Val	Glu	Val		Glu
_	~1	,	m)	245	-			1	250	_	1	~ 7		255	- 1
Pro	GIu	Val		Lys	Leu	Met	Asn		Met	Tyr	Pne	GIn	-	Asn	Ala
77.0	~1	7	260	C	~1·•	~1	17-3	265	7	T 0	C	TT i a	270	~1	7
ire	GIU	275	Pne	CÀa	сту	GIU	280	Arg	Arg	Leu	Cys	285	Ата	GIU	Arg
λκα	Lare		Dhe	Val	Sor	Glu		Tire	T Au	Tla	Thr		Gly	Lare	Dhe
Arg	290	Asp	PILE	val	ser	295	мта	тЪт	ьец	116	300	Leu	Gry	цуз	FIIE
Tle		Met	Phe	Ala	Val		Asn	Glu	Len	Lvs		Met	Lvs	Cvs	Ser
305					310		1105	014		315			_,_	0,0	320
	Lys	Asn	Asp	His		Ala	Tyr	Lys	Arq		Ala	Gln	Phe	Leu	
	•		-	325			•	•	330					335	_
Lys	Met	Ala	Asp	Pro	Gln	Ser	Ile	Gln	Glu	Ser	Gln	Asn	Leu	Ser	Met
			340					345					350		
Phe	Leu	Ala	Asn	His	Asn	Lys	Ile	Thr	Gln	Ser	Leu	Gln	Gln	Gln	Leu
		355					360					365			
Glu		Ile	Ser	Gly	Tyr		Glu	Leu	Leu	Ala		Ile	Val	Asn	Leu
	370					375	_			_	380	_			_
_	Val	Asp	Tyr	Tyr		Asn	Arg	Met	Tyr		Thr	Pro	Ser	Glu	_
385	 .	T	T	T	390		a 1	Dl	a1	395	m	T	14	D	400
HIS	Met	Leu	Leu	Lys 405	vaı	мес	GIY	Pne	410	Leu	Tyr	Leu	Met	415	GIY
Sar	V-1	Sar	λen	Ile	Tur	Larg	Τ.Δ11	λen		Lare	Luc	Δνα	Tla		T.011
561	var	Ser	420	110	TYT	цуз	neu	425	AIG	пуз	Буз	A. G	430	ASII	Deu
Ser	Lvs	Ile		Lys	Tvr	Phe	Lvs		Leu	Gln	Val	Val		Leu	Phe
	-1-	435		_,_	- 1 -		440					445			
Gly	Asp	Met	Gln	Ile	Glu	Leu	Ala	Arq	Tyr	Ile	Lys	Thr	Ser	Ala	His
•	450					455		_	=		460				
Tyr	Glu	Glu	Asn	Lys	Ser	Arg	Trp	Thr	Cys	Thr	Ser	Ser	Gly	Ser	Ser
465					470					475					480
Pro	Gln	Tyr	Asn	Ile	Cys	Glu	Gln	Met	Ile	Gln	Ile	Arg	Glu	Asp	His
				485					490					495	
Met	Arg	Phe		Ser	Glu	Leu	Ala	-	Tyr	Ser	Asn	Ser		Val	Val
			500		_	_		505					510		
Thr	Gly		Gly	Arg	Gln	Glu		Gln	Lys	Thr	Asp		Glu	Tyr	Arg
_	_	515			- 1	_	520		_		_	525	_	~1	_
Lys		Phe	Asp	Leu	Ala		Gln	Gly	Leu	Gln		Leu	Ser	Gln	Trp
0	530	772	1107	Ma =	~1	535		0		.	540	77- 7	77.2 -:	Dec	mla ac
ser	АТА	нıs	٧aı	Met	GIU	val	Tyr	ser	Trp	ьуs	ьeu	vai	HlS	Pro	ınr

545					550					55 5					560
Asp	Lys	Tyr	Ser	Asn 565	Lys	Asp	Cys	Pro	Asp 570	Ser	Ala	Glu	Glu	Tyr 575	Glu
Arg	Ala	Thr	Arg 580	Tyr	Asn	Tyr	Thr	Ser 585	Glu	Glu	Lys	Phe	Ala 590	Leu	Val
		595	Ala				600					605			
Glu	Ser 610	Val	Phe	Asn	His	Ala 615	Ile	Arg	His	Thr	Val 620	Tyr	Ala	Ala	Leu
Gln 625	Asp	Phe	Ser	Gln	Val 630	Thr	Leu	Arg	Glu	Pro 635	Leu	Arg	Gln	Ala	Ile 640
-	•	-	Lys	645					650					655	
		_	Asp 660					665					670		
	_	675	Glu		_		680		_			685			
_	690		Val	_		695					700				
705			Ser		710					715					720
			Leu	725	_				730	_				735	
_			Phe 740		_			745					750		
		75 5	Cys	_			760		_			765			
	770		Met	_		775					780				
785	_		Leu		790					795					800
		_	Val Thr	805	_				810					815	
-			820 Asn					825					830		
		835	Ala				840					845			
	850		Arg	_		855					860				
865			Asn		870					875					880
			Arg	885	-				890					895	
		-	900 Met					905					910		
		915					920					925			
	930		Leu			935					9 40				
945	_		Thr		950					955					960
	_		Met	9 65					970					975	
GTA	Arg	IIe	Thr	Leu	Hls	val	Pne	Trp	GIU	Leu	Asn	Tyr	Asp	Pne	ьeu

		985		990
Pro Asn Tyr Cys Tyr				
995 Leu Pro Phe Ser Gln		000	1005 rc Gla Pro	
	1015	III AIG ASP L	1020	Adii Ala Olii
1010 Pro Gln Tyr Leu His		ve Ala Len A		Tvr Ser Ser
1025	1030		035	1040
Ile Tyr Gly Ser Tyr				
11e lyl Gly 3cl lyl 104!		1050		1055
Ile Cys Arg Leu Leu			la Val Val	Met Glu Glu
1060	2 1	1065		1070
Leu Leu Lys Val Val	Lys Ser Le	eu Leu Gln G	ly Thr Ile	Leu Gln Tyr
1075	1	080	1085	5
Val Lys Thr Leu Met	Glu Val Me	et Pro Lys I	le Cys Arg	Leu Pro Arg
1090	1095		1100	
His Glu Tyr Gly Ser	Pro Gly I			
1105	1110		115	1120
Lys Asp Ile Val Glu			hr Val Cys	
112!		1130	T T	1135
Leu Arg Glu Val Gly	Asn Ala I.		ys Leu Leu	1150
1140	Gla Val G	1145	on Wie Ala	
Ser Leu Ser Leu Glu 1155		ys Asp Led L 160	eu nis Aia 1165	
Gln Asn Ile Leu Pro				
1170	1175	.15 var 2/2 c	1180	-
Ala Lys Met Lys Arg		er Lys Tyr A		His Leu Val
1185	1190		195	1200
Pro Leu Ile Glu Arg	Leu Gly T	hr Pro Gln G	ln Ile Ala	Ile Ala Arg
120		1210		1215
Glu Gly Asp Leu Leu			ys Cys Gly	Leu Ser Met
Glu Gly Asp Leu Leu 1220	Thr Lys G	lu Arg Leu C 1225		Leu Ser Met 1230
Glu Gly Asp Leu Leu	Thr Lys G	lu Arg Leu C 1225 le Arg Ser P	he Leu Asp	Leu Ser Met 1230 Asp Pro Ile
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235	Thr Lys G Thr Arg I	lu Arg Leu C 1225 le Arg Ser P 240	he Leu Asp 124!	Leu Ser Met 1230 Asp Pro Ile
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu	Thr Lys G Thr Arg I Pro Ser A	lu Arg Leu C 1225 le Arg Ser P 240	he Leu Asp 124! et His Val	Leu Ser Met 1230 Asp Pro Ile
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250	Thr Lys G Thr Arg I 1 Pro Ser A 1255	lu Arg Leu C 1225 le Arg Ser P 240 sn Gly Val M	he Leu Asp 124! et His Val 1260	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S	Elu Arg Leu C 1225 Tle Arg Ser P 240 Len Gly Val M	he Leu Asp 124! et His Val 1260 In Phe Val	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270	Elu Arg Leu C 1225 Tle Arg Ser P 240 Isn Gly Val M Eer Ala Met G	he Leu Asp 124! et His Val 1260 ln Phe Val 275	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T	Elu Arg Leu C 1225 Ele Arg Ser P 240 Isn Gly Val M Fer Ala Met G 1 Chr Val Glu G	he Leu Asp 124! et His Val 1260 ln Phe Val 275	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T	Elu Arg Leu C 1225 Ele Arg Ser P 240 Isn Gly Val M Eer Ala Met G 1 Thr Val Glu G 1290	he Leu Asp 124! et His Val 1260 ln Phe Val 275 ln Cys Phe	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T	Elu Arg Leu C 1225 Ele Arg Ser P 240 Isn Gly Val M Eer Ala Met G 1 Thr Val Glu G 1290	he Leu Asp 124! et His Val 1260 ln Phe Val 275 ln Cys Phe	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly 1300	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T Cys Met I	Elu Arg Leu C 1225 Ele Arg Ser P 240 Isn Gly Val M Eer Ala Met G 1 Chr Val Glu G 1290 Ele Ile Val L 1305	he Leu Asp 124! et His Val 1260 ln Phe Val 275 ln Cys Phe	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295 Gln Gln Arg 1310
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T Cys Met I Asp Phe C	Elu Arg Leu C 1225 Ele Arg Ser P 240 Isn Gly Val M Eer Ala Met G 1 Chr Val Glu G 1290 Ele Ile Val L 1305	he Leu Asp 124! et His Val 1260 ln Phe Val 275 ln Cys Phe	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295 Gln Gln Arg 1310 Val Gln Lys
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly 1300 Arg Phe Ala Val Leu 1315	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T Cys Met I Asp Phe C	Elu Arg Leu C 1225 Tle Arg Ser P 240 Asn Gly Val M Ger Ala Met G 1 Thr Val Glu G 1290 Tle Ile Val L 1305 Tys Tyr His L	he Leu Asp 124! et His Val 1260 ln Phe Val 275 ln Cys Phe eu Leu Gly eu Leu Lys 132!	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295 Gln Gln Arg 1310 Val Gln Lys
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly 1300 Arg Phe Ala Val Leu 1315 His Asp Gly Lys Asp	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T Cys Met I Asp Phe C Glu Ile I 1335	elu Arg Leu C 1225 The Arg Ser P 240 The Ala Met G Thr Val Glu G 1290 The Ile Val L 1305 Tys Tyr His L 320 The Lys Asn V	he Leu Asp 124! et His Val 1260 ln Phe Val 275 ln Cys Phe eu Leu Gly eu Leu Lys 132! al Pro Leu 1340	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295 Gln Gln Arg 1310 Val Gln Lys 5 Lys Lys Met
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly 1300 Arg Phe Ala Val Leu 1315 His Asp Gly Lys Asp	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T Cys Met I Asp Phe C Glu Ile I 1335	elu Arg Leu C 1225 The Arg Ser P 240 The Ala Met G Thr Val Glu G 1290 The Ile Val L 1305 Tys Tyr His L 320 The Lys Asn V	he Leu Asp 124! et His Val 1260 ln Phe Val 275 ln Cys Phe eu Leu Gly eu Leu Lys 132! al Pro Leu 1340	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295 Gln Gln Arg 1310 Val Gln Lys 5 Lys Lys Met Ile Ile Thr
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly 1300 Arg Phe Ala Val Leu 1315 His Asp Gly Lys Asp 1330 Val Glu Arg Ile Arg	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T Cys Met I Asp Phe C Glu Ile I 1335 Lys Phe G	Elu Arg Leu C 1225 Ele Arg Ser P 240 Esn Gly Val M Eer Ala Met G Thr Val Glu G 1290 Ele Ile Val L 1305 Eys Tyr His L 320 Ele Lys Asn V Eln Ile Leu A	he Leu Asp 1249 et His Val 1260 In Phe Val 275 In Cys Phe eu Leu Gly eu Leu Lys 1329 al Pro Leu 1340 sn Asp Glu 355	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295 Gln Gln Arg 1310 Val Gln Lys 5 Lys Lys Met Ile Ile Thr 1360
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly 1300 Arg Phe Ala Val Leu 1315 His Asp Gly Lys Asp 1330 Val Glu Arg Ile Arg 1345 Ile Leu Asp Lys Tyr	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T Cys Met I Asp Phe C Glu Ile I 1335 Lys Phe G 1350 Leu Lys S	Plu Arg Leu C 1225 The Arg Ser P 240 The Arg Ser P 240 The Ala Met G Thr Val Glu G 1290 The Ile Val L 1305 Tys Tyr His L 320 The Lys Asn V The Leu A The Coly Asp G	he Leu Asp 1249 et His Val 1260 In Phe Val 275 In Cys Phe eu Leu Gly eu Leu Lys 1329 al Pro Leu 1340 sn Asp Glu 355	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295 Gln Gln Arg 1310 Val Gln Lys 5 Lys Lys Met Ile Ile Thr 1360 Thr Pro Val
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly 1300 Arg Phe Ala Val Leu 1315 His Asp Gly Lys Asp 1330 Val Glu Arg Ile Arg 1345 Ile Leu Asp Lys Tyr 136	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T Cys Met I Asp Phe C Glu Ile I 1335 Lys Phe G 1350 Leu Lys S	Elu Arg Leu C 1225 Ele Arg Ser P 240 Esn Gly Val M Eer Ala Met G Thr Val Glu G 1290 Ele Ile Val L 1305 Eys Tyr His L 320 Ele Lys Asn V Eln Ile Leu A Eer Gly Asp G 1370	he Leu Asp 124! et His Val 1260 In Phe Val 275 In Cys Phe eu Leu Gly eu Leu Lys 132! al Pro Leu 1340 sn Asp Glu 355 ly Glu Gly	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295 Gln Gln Arg 1310 Val Gln Lys 5 Lys Lys Met Ile Ile Thr 1360 Thr Pro Val 1375
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly 1300 Arg Phe Ala Val Leu 1315 His Asp Gly Lys Asp 1330 Val Glu Arg Ile Arg 1345 Ile Leu Asp Lys Tyr 136 Glu His Val Arg Cys	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T Cys Met I Asp Phe C Glu Ile I 1335 Lys Phe G 1350 Leu Lys S	Plu Arg Leu C 1225 The Arg Ser P 240 The Arg Ser P 240 The Ala Met G Thr Val Glu G 1290 The Ile Val L 1305 Tys Tyr His L 320 The Lys Asn V The Gly Asp G 1370 The Pro Ile H	he Leu Asp 124! et His Val 1260 In Phe Val 275 In Cys Phe eu Leu Gly eu Leu Lys 132! al Pro Leu 1340 sn Asp Glu 355 ly Glu Gly	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295 Gln Gln Arg 1310 Val Gln Lys 5 Lys Lys Met Ile Ile Thr 1360 Thr Pro Val 1375 Leu Ala Ser
Glu Gly Asp Leu Leu 1220 Phe Glu Val Ile Leu 1235 Trp Arg Gly Pro Leu 1250 Val Glu Phe His Arg 1265 Pro Val Gly Thr His 128 Leu His Trp Ala Gly 1300 Arg Phe Ala Val Leu 1315 His Asp Gly Lys Asp 1330 Val Glu Arg Ile Arg 1345 Ile Leu Asp Lys Tyr 136	Thr Lys G Thr Arg I Pro Ser A 1255 Leu Trp S 1270 Glu Phe T Cys Met I Asp Phe C Glu Ile I 1335 Lys Phe G 1350 Leu Lys S	Elu Arg Leu C 1225 Ele Arg Ser P 240 Esn Gly Val M Eer Ala Met G Thr Val Glu G 1290 Ele Ile Val L 1305 Eys Tyr His L 320 Ele Lys Asn V Eln Ile Leu A Eer Gly Asp G 1370	he Leu Asp 124! et His Val 1260 In Phe Val 275 In Cys Phe eu Leu Gly eu Leu Lys 132! al Pro Leu 1340 sn Asp Glu 355 ly Glu Gly	Leu Ser Met 1230 Asp Pro Ile 5 Asp Glu Cys Tyr Cys Ile 1280 Gly Asp Gly 1295 Gln Gln Arg 1310 Val Gln Lys 5 Lys Lys Met Ile Ile Thr 1360 Thr Pro Val 1375

```
<210> 177
<211> 417
<212> DNA
<213> Homo sapiens
<400> 177
acgegtgatg teacactgee tetgeegetg ggteetaatt egattgeacg caccatgget
gcagttcgtg gcgcgcatag tttctggcat gcttcgcgca tcctggagac cgatcccgcc
getgeegtga aacegeetaa aaatgtgaag egattgeeca aageegtgte egtggageaa
atgcaaaagc teettgecat acceagtett aagaeteeta eeggeetgeg taategageg
atacttgagt tettatatge taceggegeg egegtgageg agatgetgge aacagacetg
gacgatatac acctgggcga aaaaccccgc gatgaaaacg gggaatctat tgcacttccc
gggtatgtgc gcctttttgg aaagggaggt aaagagcgtt tagtcccttt gggatcc
417
<210> 178
<211> 139
<212> PRT
<213> Homo sapiens
<400> 178
Thr Arg Asp Val Thr Leu Pro Leu Pro Leu Gly Pro Asn Ser Ile Ala
                 5
Arg Thr Met Ala Ala Val Arg Gly Ala His Ser Phe Trp His Ala Ser
            20
                                25
Arg Ile Leu Glu Thr Asp Pro Ala Ala Ala Val Lys Pro Pro Lys Asn
                            40
Val Lys Arg Leu Pro Lys Ala Val Ser Val Glu Gln Met Gln Lys Leu
Leu Ala Ile Pro Ser Leu Lys Thr Pro Thr Gly Leu Arg Asn Arg Ala
                                        75
                    70
Ile Leu Glu Phe Leu Tyr Ala Thr Gly Ala Arg Val Ser Glu Met Leu
                                    90
Ala Thr Asp Leu Asp Asp Ile His Leu Gly Glu Lys Pro Arg Asp Glu
                                105
            100
Asn Gly Glu Ser Ile Ala Leu Pro Gly Tyr Val Arg Leu Phe Gly Lys
                            120
Gly Gly Lys Glu Arg Leu Val Pro Leu Gly Ser
                        135
   1.30
<210> 179
<211> 362
<212> DNA
<213> Homo sapiens
<400> 179
acgegtegaa ggtgeeggtg ggggegatea ataacatege geaateeetg gaagageete
```

aggtgattgc ccgtgggttg atggtggaag atcccgcatc cccaagaatc cgggaattcg

```
120
ccattgggcc gggcagcccg aatccaaaat gtcggggcac gcccagtggg agtatggtaa
ggggccggca ccgatgttgg nggcagcata cggatggaag tgctgggcga gcgcctgggt
240
ttgccggcag agcaactggg gcagctcaag gcgggcgggg tgatcgagca gttggattga
gcaatggcgg ccgcgaagcc cgccatttac cttgatgact gtttagcgcg cggattcttt
360
aa
362
<210> 180
<211> 108
<212> PRT
<213> Homo sapiens
<400> 180
Met Ala Gly Phe Ala Ala Ala Ile Ala Gln Ser Asn Cys Ser Ile Thr
                                    10
Pro Pro Ala Leu Ser Cys Pro Ser Cys Ser Ala Gly Lys Pro Arg Arg
                                25
Ser Pro Ser Thr Ser Ile Arg Met Leu Pro Pro Thr Ser Val Pro Ala
                            40
Pro Tyr His Thr Pro Thr Gly Arg Ala Pro Thr Phe Trp Ile Arg Ala
                        55
Ala Arg Pro Asn Gly Glu Phe Pro Asp Ser Trp Gly Cys Gly Ile Phe
                                        75
                    70
His His Gln Pro Thr Gly Asn His Leu Arg Leu Phe Gln Gly Leu Arg
                                    90
Asp Val Ile Asp Arg Pro His Arg His Leu Arg Arg
            100
                                105
<210> 181
<211> 297
<212> DNA
<213> Homo sapiens
<400> 181
gegttgatea tgteegacee aggettgate atgetggtae geegteactt ecegtgeatg
ccgattcact tgtcggtaca ggccaatacg gtgaattggg ccagcgtcga gttctggcaa
cagcaaggta tetgeegggt aatcetgteg egggaattgt caetggaaga aateggegaa
atccgccaac aggtgccggc catggagctg gaagtgtttg tgcacggtgc cctgtacatg
qcctattccq qqcqctqttt gttgtccggc tatatgaaca agcgcgatgc caaccaa
297
<210> 182
<211> 99
<212> PRT
```

<213> Homo sapiens <400> 182 Ala Leu Ile Met Ser Asp Pro Gly Leu Ile Met Leu Val Arg Arg His 5 Phe Pro Cys Met Pro Ile His Leu Ser Val Gln Ala Asn Thr Val Asn Trp Ala Ser Val Glu Phe Trp Gln Gln Gln Gly Ile Cys Arg Val Ile 40 Leu Ser Arg Glu Leu Ser Leu Glu Glu Ile Gly Glu Ile Arg Gln Gln 55 Val Pro Ala Met Glu Leu Glu Val Phe Val His Gly Ala Leu Tyr Met 70 75 Ala Tyr Ser Gly Arg Cys Leu Leu Ser Gly Tyr Met Asn Lys Arg Asp 90 Ala Asn Gln <210> 183 <211> 351 <212> DNA <213> Homo sapiens <400> 183 cgggacgtca ccatgaagcc gaccggctcg ggggatgtgg cgaacaaggt catcacccat atteegttta acategtete ecaggegaet catecattee ttegtaeett ggaegatgte aagcgcatct ctttggcgac cgacgggctc ggccaccagg tcctgctcaa gggctaccag geogagggee acquetaege acacecegae taeggeggea acgteteeca cegtgeegge gggatgaagg atctcgagaa gctcaccgag tegggcaggc agtggaacac cgatttcggc attcacgtca acctggtgga gtcctatcct gaggcgaatc acttcggcga c 351 <210> 184 <211> 117 <212> PRT <213> Homo sapiens <400> 184 Arg Asp Val Thr Met Lys Pro Thr Gly Ser Gly Asp Val Ala Asn Lys Val Ile Thr His Ile Pro Phe Asn Ile Val Ser Gln Ala Thr His Pro 20 25 Phe Leu Arg Thr Leu Asp Asp Val Lys Arg Ile Ser Leu Ala Thr Asp Gly Leu Gly His Gln Val Leu Leu Lys Gly Tyr Gln Ala Glu Gly His 60 Asp Tyr Ala His Pro Asp Tyr Gly Gly Asn Val Ser His Arg Ala Gly Gly Met Lys Asp Leu Glu Lys Leu Thr Glu Ser Gly Arg Gln Trp Asn

```
85
                                    90
Thr Asp Phe Gly Ile His Val Asn Leu Val Glu Ser Tyr Pro Glu Ala
            100
                               105
Asn His Phe Gly Asp
        115
<210> 185
<211> 396
<212> DNA
<213> Homo sapiens
<400> 185
egegtgggte teagtaaaga aaatttgttg ettagaggat geaceattag aaacacagag
gctgttgtgg gcattgtggt ttatgcaggc catgaaacca aagcaatgct gaacaacagt
gggccacggt ataagcgcag caaattagaa agaagagcaa acacagatgt cctctggtgt
180
qtcatqcttc tqqtcataat qtqcttaact qqcqcaqtaq qtcatqqaat ctqqctqaqc
aggtatgaaa agatgcattt tttcaatgtt cccgagcctg atggacatat catatcacca
300
etgttggcag gattttatat gttttggace gtgatcattt tgttacaggt ettgatteet
atttctctct atgtttccat cgaaattgtg aagctt
396
<210> 186
<211> 132
<212> PRT
<213> Homo sapiens
<400> 186
Arg Val Gly Leu Ser Lys Glu Asn Leu Leu Arg Gly Cys Thr Ile
1
                 5
                                    10
Arg Asn Thr Glu Ala Val Val Gly Ile Val Val Tyr Ala Gly His Glu
Thr Lys Ala Met Leu Asn Asn Ser Gly Pro Arg Tyr Lys Arg Ser Lys
                            40
Leu Glu Arg Arg Ala Asn Thr Asp Val Leu Trp Cys Val Met Leu Leu
                       55
                                            60
Val Ile Met Cys Leu Thr Gly Ala Val Gly His Gly Ile Trp Leu Ser
                    70
                                        75
Arg Tyr Glu Lys Met His Phe Phe Asn Val Pro Glu Pro Asp Gly His
                                   90
Ile Ile Ser Pro Leu Leu Ala Gly Phe Tyr Met Phe Trp Thr Val Ile
           100
                               105
                                                   110
Ile Leu Leu Gln Val Leu Ile Pro Ile Ser Leu Tyr Val Ser Ile Glu
       115
                          120
                                               125
Ile Val Lys Leu
   130
<210> 187
<211> 423
```

```
<212> DNA
<213> Homo sapiens
<400> 187
egagtgetea eegegteage egteatgegt eecaetgagg etgttgtete teggteggea
gaacetegae gagtteageg gateetggae eagegegagt gggetggegt ettegttgte
gatgagcatc gtcgtttgct tggcacggtc ggcgatcaag aggtcatcga ggctgctcgc
egeggagate geagtattge tgaegeggtg gaaactaaeg geateeteae ggegeggaee
gacactecgt tgteegaget ettegeteeg accageaacg ceagggtgee gttggeegtt
300
gtegacgagg acttecacet catgggtgte ateteteggg tgaccetget egacgegatg
tcacgagctc gcgacgaggc aggagaggga tctgtcatgt ccttggagaa caccggaaag
420
ctt
423
<210> 188
<211> 141
<212> PRT
<213> Homo sapiens
<400> 188
Arg Val Leu Thr Ala Ser Ala Val Met Arg Pro Thr Glu Ala Val Val
                                    10
Ser Arg Ser Ala Glu Pro Arg Arg Val Gln Arg Ile Leu Asp Gln Arg
                                25
Glu Trp Ala Gly Val Phe Val Val Asp Glu His Arg Arg Leu Leu Gly
                            40
Thr Val Gly Asp Gln Glu Val Ile Glu Ala Ala Arg Arg Gly Asp Arg
Ser Ile Ala Asp Ala Val Glu Thr Asn Gly Ile Leu Thr Ala Arg Thr
65
                                        75
                    70
Asp Thr Pro Leu Ser Glu Leu Phe Ala Pro Thr Ser Asn Ala Arg Val
                                    90
Pro Leu Ala Val Val Asp Glu Asp Phe His Leu Met Gly Val Ile Ser
                                105
Arg Val Thr Leu Leu Asp Ala Met Ser Arg Ala Arg Asp Glu Ala Gly
        115
                            120
Glu Gly Ser Val Met Ser Leu Glu Asn Thr Gly Lys Leu
                        135
<210> 189
<211> 429
<212> DNA
<213> Homo sapiens
<400> 189
ngatggttta ccaacatatg cacggttcga gcggcaatag ctcctcgggg gctggcagtg
```

```
aaatgtttga agatgccggc gtttccggcc tcaacttgtt tcgatgccgt ggttccaccg
120
atttegeega tgeggeteat egeaeggtta agaagttteg teeagataac ccaggacaga
180
gcaaggtata tcaggctcag aaccaggaaa agcagggctt taccccagtg ccccatatag
accgcgctag ctacggcaaa aggcgcgccc agtggggtcc aggacagcac tttcatggct
gaagggageg catecenage ttegectage eccagageta acceagegae cagtggacea
gegeceatea teagtaggaa eectaegata ateageeett gttttaeeee tggaatggag
420
ctgatttcn
429
<210> 190
<211> 123
<212> PRT
<213> Homo sapiens
<400> 190
Met Met Gly Ala Gly Pro Leu Val Ala Gly Leu Ala Leu Gly Leu Gly
                                    10
Glu Ala Xaa Asp Ala Leu Pro Ser Ala Met Lys Val Leu Ser Trp Thr
                                25
Pro Leu Gly Ala Pro Phe Ala Val Ala Ser Ala Val Tyr Met Gly His
                            40
Trp Gly Lys Ala Leu Leu Phe Leu Val Leu Ser Leu Ile Tyr Leu Ala
                        55
Leu Ser Trp Val Ile Trp Thr Lys Leu Leu Asn Arg Ala Met Ser Arg
                    70
                                        75
Ile Gly Glu Ile Gly Gly Thr Thr Ala Ser Lys Gln Val Glu Ala Gly
                                    90
Asn Ala Gly Ile Phe Lys His Phe Thr Ala Ser Pro Arg Gly Ala Ile
                                105
Ala Ala Arg Thr Val His Met Leu Val Asn His
        115
<210> 191
<211> 4845
<212> DNA
<213> Homo sapiens
<400> 191
ccgcccgggg ccatggcgac actcagcttc gtcttcctgc tgctgggggc agtgtcctgg
cotteeggett etgeeteegg ceaggagtte tggeeeggae aateggegge egatattetg
tegggggegg etteeegeag aeggtatett etgtatgaeg teaaceeece ggaaggette
aacctgcgca gggatgtcta tatccgaatc gcctctctcc tgaagactct gctgaagacg
gaggagtggg tgcttgtcct gcctccatgg ggccgcctct atcactggca gagtcctgac
```

atccaccagg	tccggattcc	ctggtctgag	ttttttgatc	ttccaagtct	caataaaaac
atccccgtca 420	tcgagtatga	gcagttcatc	gcagaatctg	gtgggccctt	tattgaccag
480			tggaaagaag		
540			tactcccagg		
600			ggtctaaacg		
660			ctgagaaaca		
720			cactatggag		
780			cgggaggtgg		
840			atccccttcc		
900			ccctacctgg		
960			cccagtctgg		
1020			aaggtgtttg		
1080			cccgagatgg		
1140			gttgcgatta		
1200			gccgagagcg cagatggtcg		
1260					
1320			teetgtagea		
1380			tagttctcat		
1440			aacctcttct		
1500			ccctgagagt		
1560			ccgtgagcgt		
1620			catgagcgtc		
1680			gtgagcgtcc		
1740			tgagcgtccc		
1800			tgggaggtac		
1860			tgaaggetgg		
agttaaatgg 1920	aaaatgaaag	tagagggaat	gatetteece	gtggttagca	ctgtgcacac

gcgtgcgtct 1980	ctgtgggtta	gtctgtctct	ctcctgcccg	aggaatgctg	agcgccctga
	cttcacacat	ctgctatttc	ctgtggtgtt	ctgggcatgg	tgtataagac
	tccgggtgat	gctgtctgct	gggtgtgggt	ccctttccct	gttaagcaga
	cgctgacttc	ttaggtcagg	gcggaggtgt	gcaggagccc	agtcacgagc
	ttctcaggtg	tggccttggg	attttgactg	cgaccttggc	ggtgctgtct
	ggaagcctgc	tgtggggagg	cttcgcactg	agctctcagc	ctcctgccct
	aagcgctcgg	cccagctcac	tgaagctgcc	ctgcctccgg	gccggcgcgg
	caggcccctg	tgtgtggggt	ggtgagggtc	tececaceag	tgctgcaccc
cgcagcagca 2460	tacaggcctg	tgtggcctgc	tggccctgtg	gctctgtgta	cagcgctgtg
catgttacat 2520	ttgctctgga	aacatctctg	gggtttgctt	gttcacgaag	ttcatgaagt
gccgctggag 2580	agccagagac	cagctgcgca	ggagccggag	gaacgggcag	geegetgace
tgaggtctgg 2640	agaaacccct	ggagaagggt	gtccccacca	gcccatacag	cgtgtgtgtg
gagggggcct 2700	tgacctccgt	gatgtctact	gtgcctcagg	ataaggaccc	gccatgccct
ggctagacag 2760	tgtgtggtta	gtaggaatct	ctcattgttc	accatgtgac	ccccaggagg
2820				acggagagga	
2880				gtggggtgtg	
2940				cacctgagcg	
3000				gtgttggttc	
3060				gactctggaa	
3120				aggtttttt	
3180				aatcctgggg	
3240				gtgtgagcaa	
3300				ccggacccga	
3360				cagccgggag	
3420				cacgcctcta	
3480				ctccgctgct	
geggttgttt 3540	tcaggcagcg	tetgtgaace	cacagetegg	ttgccagcag	tgcccgcgtg

```
gtgacccaga agcaggagtg tttgtcaggc tcccgctctg gcctttccag ccacctttca
tgtcttcata ttttaagtgc attgaggata gatgcaggcg ggtgagctgc cctccgtcag
3660
gtggaccegg getgacattt eeetgggage tggtgeaagg agaagegtea ttttaaatgt
ctgcagageg accaggggee teatgaatet eteegttgee eteegegeag eaggaggetg
3780
cetgtgtgtt teeteetggg aegegtgeaa ggeagaeetg gtgetgeaaa ggaaagggee
tgaggcctca gggagccccg tggagggatg acagttcagg ccctactgct ggcacgtcag
3900
ageactggga agtititeag tgacgietet ggggeactea giggatigie igiaggaaae
ttgcagetet geteeteaca ecaggeeegg etggeeacce accetegeee ecaetggeea
cccctccctc gccccgactg ccccgcccca ccctcacccc gactgccccg ccctcgcccg
getggeegte cetgeecteg ceeeggetgg caggtgeaca tggggeetee aggtetgeea
4140
ttcgctattg agaactagaa atgaggaagg acagttacgc taactccaaa aggctgtcta
ggatgagetg ctttateagg gageteettg tacceatttt acagaaatca tttttaggte
tttgtgccac caccacgagg ggcatctgca aagagggcaa cgctagacac agaatccgtg
gaaggtgcag cagtgcctca ggggtcctca gggtcaggga gcccccctca ccctcttggc
cogttaccet ttgtgacttt ccaccatggt gtcgtgtgac cctcagtcag gttggtgggg
getgagteet caetgageag ceaettteea catetgetag aggaacagtg acatggacae
ctgtgacaga gagaggacag ttagtgagga gggacagaca gctcttcctt tcggagcctg
gctagtctag gacatcacct tgctgtgtct tctcaagctt ttaaaattga ccctgaacgt
4620cctatggtgt tactcaaagc tgtgcagggt aaatgatgac atatttattc
tttttccatt tgttctagaa acagtgcctt tttcatcagt tgcattttcc aggctgagag
4740
ctgtataaaa cattttggac tgtgaccatg taccttcctt tttaagaaaa ataaactgct
4800
4845
<210> 192
<211> 428
<212> PRT
<213> Homo sapiens
<400> 192
Pro Pro Gly Ala Met Ala Thr Leu Ser Phe Val Phe Leu Leu Gly
                                   10
Ala Val Ser Trp Pro Pro Ala Ser Ala Ser Gly Gln Glu Phe Trp Pro
Gly Gln Ser Ala Ala Asp Ile Leu Ser Gly Ala Ala Ser Arg Arg Arg
```

```
40
Tyr Leu Leu Tyr Asp Val Asn Pro Pro Glu Gly Phe Asn Leu Arg Arg
                  55
Asp Val Tyr Ile Arg Ile Ala Ser Leu Leu Lys Thr Leu Leu Lys Thr
                    75
      70
Glu Glu Trp Val Leu Val Leu Pro Pro Trp Gly Arg Leu Tyr His Trp
                   90
Gln Ser Pro Asp Ile His Gln Val Arg Ile Pro Trp Ser Glu Phe Phe
       100 105 110
Asp Leu Pro Ser Leu Asn Lys Asn Ile Pro Val Ile Glu Tyr Glu Gln
           120
Phe Ile Ala Glu Ser Gly Gly Pro Phe Ile Asp Gln Val Tyr Val Leu
                                    140
                   135
Gln Ser Tyr Ala Glu Gly Trp Lys Glu Gly Thr Trp Glu Glu Lys Val
                                155
               150
Asp Glu Arg Pro Cys Ile Asp Gln Leu Leu Tyr Ser Gln Asp Lys His
                             170 175
            165
Glu Tyr Tyr Arg Gly Trp Phe Trp Gly Tyr Glu Glu Thr Arg Gly Leu
                         185
Asn Val Ser Cys Leu Ser Val Gln Gly Ser Ala Ser Ile Val Ala Pro
                       200
Leu Leu Leu Arg Asn Thr Ser Ala Arg Ser Val Met Leu Asp Arg Ala
                   215
Glu Asn Leu Leu His Asp His Tyr Gly Gly Lys Glu Tyr Trp Asp Thr
                230
                                 235
Arg Arg Ser Met Val Phe Ala Arg His Leu Arg Glu Val Gly Asp Glu
                             250 255
       245
Phe Arg Ser Arg His Leu Asn Ser Thr Asp Asp Ala Asp Arg Ile Pro
         260
                         265
Phe Gln Glu Asp Trp Met Lys Met Lys Val Lys Leu Gly Ser Ala Leu
                      280
Gly Gly Pro Tyr Leu Gly Val His Leu Arg Arg Lys Asp Phe Ile Trp
   290 295
                                    300
Gly His Arg Gln Asp Val Pro Ser Leu Glu Gly Ala Val Arg Lys Ile
   310
                                315 320
Arg Ser Leu Met Lys Thr His Arg Leu Asp Lys Val Phe Val Ala Thr
                            330 335
Asp Ala Val Arg Lys Glu Tyr Glu Glu Leu Lys Lys Leu Leu Pro Glu
                         345
         340
Met Val Arq Phe Glu Pro Thr Trp Glu Glu Leu Glu Leu Tyr Lys Asp
                       360
Gly Gly Val Ala Ile Ile Asp Gln Trp Ile Cys Ala His Ala Arg Cys
                   375
Leu Pro Thr Ser Leu Ser Ala Glu Ser Gly Ser Gly Gly Phe Gln Arg
                                395
      390
Phe Phe Cys Pro Lys Tyr Ser Val Ser Glu Gln Met Val Ala Cys Val
                   410
            405
His Ser Gly His Phe His Thr Val Cys Leu Leu Val
         420
                          425
<210> 193
<211> 350
<212> DNA
```

<213> Homo sapiens

```
<400> 193
geeggegage tggaetgege cateatggee gagecettee eegacaeegg eetggeeaeg
gegeagetgt acgaegagee ettegtegte gegetgeggg egtegeacee getggeegae
cgtgccagca tcagccccga ggaggtcaag ggcgagacca tgttgatgtt gggcacgggc
180
ccctggtttc cccgggcccg cggtgggggt ttggcccgga tttggcgcgt ttctccagcg
cegttaaggg catacgccgc agtttcgagg gctcgtcgct ggagaccatc aagcacatcg
tggcttcggg catggcgtga cggtggtgcc gcagctgtcc gtgccgcgcg
350
<210> 194
<211> 116
<212> PRT
<213> Homo sapiens
<400> 194
Ala Gly Glu Leu Asp Cys Ala Ile Met Ala Glu Pro Phe Pro Asp Thr
                                    10
Gly Leu Ala Thr Ala Gln Leu Tyr Asp Glu Pro Phe Val Val Ala Leu
                                25
Arg Ala Ser His Pro Leu Ala Asp Arg Ala Ser Ile Ser Pro Glu Glu
                            40
Val Lys Gly Glu Thr Met Leu Met Leu Gly Thr Gly Pro Trp Phe Pro
Arg Ala Arg Gly Gly Leu Ala Arg Ile Trp Arg Val Ser Pro Ala
                    70
Pro Leu Arg Ala Tyr Ala Ala Val Ser Arg Ala Arg Arg Trp Arg Pro
                                    90
Ser Ser Thr Ser Trp Leu Arg Ala Trp Arg Asp Gly Gly Ala Ala Ala
            100
Val Arg Ala Ala
        115
<210> 195
<211> 495
<212> DNA
<213> Homo sapiens
<400> 195
acgcgtgaac gcgacggctt ggcgatcgga ggcgtcggcc ccgtcgttga gtgggccgtt
gaaatggttc gcttcgacga aagcgagact ctcgaccgcc ttgcatcggg cgtccttgaa
ccagaacttg gcgacgattt ggccgccgtc ctgctcgatt ctcatcgggt tgctgtcatc
agegagggat cgaactggct tgcctcgcta cccgtgatcg taggtcgcaa cacggaacag
tttcgcagca taccagacct tgcccgcgac cggatcgaca aactgcacca gttgagccat
300
```

```
egegaaatag caegaaateg egageteetg egtgeeegeg etgegteggg geaggtgegg
cactqccacq qcqacqcaca cctcqqcaac atcqtcatqa ttqacqqcaa gccqqtcctq
ttegacqega tegaatttga teetgatate gegacaaegg atgtgetgta egatttegeg
ttccctctga tggat
495
<210> 196
<211> 165
<212> PRT
<213> Homo sapiens
<400> 196
Thr Arg Glu Arg Asp Gly Leu Ala Ile Gly Gly Val Gly Pro Val Val
                                    10
Glu Trp Ala Val Glu Met Val Arg Phe Asp Glu Ser Glu Thr Leu Asp
                                25
Arg Leu Ala Ser Gly Val Leu Glu Pro Glu Leu Gly Asp Asp Leu Ala
                            40
Ala Val Leu Leu Asp Ser His Arg Val Ala Val Ile Ser Glu Gly Ser
                        55
                                            60
Asn Trp Leu Ala Ser Leu Pro Val Ile Val Gly Arg Asn Thr Glu Gln
                                        75
                    70
Phe Arg Ser Ile Pro Asp Leu Ala Arg Asp Arg Ile Asp Lys Leu His
                85
                                    90
Gln Leu Ser His Arg Glu Ile Ala Arg Asn Arg Glu Leu Leu Arg Ala
                                105
Arg Ala Ala Ser Gly Gln Val Arg His Cys His Gly Asp Ala His Leu
                            120
        115
Gly Asn Ile Val Met Ile Asp Gly Lys Pro Val Leu Phe Asp Ala Ile
Glu Phe Asp Pro Asp Ile Ala Thr Thr Asp Val Leu Tyr Asp Phe Ala
145
                    150
                                        155
Phe Pro Leu Met Asp
                165
<210> 197
<211> 402
<212> DNA
<213> Homo sapiens
<400> 197
caagcaatgc ttgacgcagt tgttgaatac ttaccagcac cgactgatat tccagcaatc
aaaggtatca atccagatga aactgaaggt qaacgtcacg caagcgatga tgagccattc
tetteattag catteaaaat tgeaactgae ceattegtag gtaacttaae ettetteegt
gtgtactcag gtgtaattaa ctctggtgat acagtattaa actctgtacg tcaaaaacgt
240
gaacgttttg gtcgtatcgt acagatgcac gctaataaac gtgaagaaat taaagaagtt
```

```
cgtgcgggcg atatcgctgc agcaatcggc ttaaaagatg taactacggg tgaaccatta
tgtgctgtcg atgcaccaat cattcttgag cgtatggaat tc
402
<210> 198
<211> 134
<212> PRT
<213> Homo sapiens
<400> 198
Gln Ala Met Leu Asp Ala Val Val Glu Tyr Leu Pro Ala Pro Thr Asp
                                    10
Ile Pro Ala Ile Lys Gly Ile Asn Pro Asp Glu Thr Glu Gly Glu Arg
                                25
His Ala Ser Asp Asp Glu Pro Phe Ser Ser Leu Ala Phe Lys Ile Ala
                            40
Thr Asp Pro Phe Val Gly Asn Leu Thr Phe Phe Arg Val Tyr Ser Gly
Val Ile Asn Ser Gly Asp Thr Val Leu Asn Ser Val Arg Gln Lys Arg
                    70
                                        75
Glu Arg Phe Gly Arg Ile Val Gln Met His Ala Asn Lys Arg Glu Glu
                85
                                    90
Ile Lys Glu Val Arg Ala Gly Asp Ile Ala Ala Ile Gly Leu Lys
                               105
            100
Asp Val Thr Thr Gly Glu Pro Leu Cys Ala Val Asp Ala Pro Ile Ile
        115
                            120
                                                125
Leu Glu Arg Met Glu Phe
    130
<210> 199
<211> 507
<212> DNA
<213> Homo sapiens
<400> 199
acgcgtgaag tcgtgcatag atcggtgtga catagagaag cctccgaccc aagctgcgta
tatoqoacaa agaccaagog accotggacg ttotagacag aactotgota cgaggootga
120
caatagtgaa atccccgaga acccagctat ggaagggttt ccagatgctc gaaggcctgt
cataccagag gttaggttaa actgtatgga gactttcgag gtgaaagttg actcgccggt
aaagcctgct cctaaagagg atttagatct gatagatcta tcctcagatt caacctcggg
geetgaaaaa caetetatae teteaacete egacagegae tetettgtat ttgageetet
tecetetete agaatagteg agagtgaega agaagaggag aegatgaaee aaggegatga
eggecetee ggtaaaaatg etgeetette teectecate eccagecate ceteegteet
cageetgage acageteege ttgtaca
507
```

```
<210> 200
<211> 153
<212> PRT
<213> Homo sapiens
<400> 200
Met Glu Gly Glu Glu Ala Ala Phe Leu Pro Glu Gly Pro Ser Ser Pro
Trp Phe Ile Val Ser Ser Ser Ser Leu Ser Thr Ile Leu Arg Glu
                                25
Gly Arg Gly Ser Asn Thr Arg Glu Ser Leu Ser Glu Val Glu Ser Ile
                            40
Glu Cys Phe Ser Gly Pro Glu Val Glu Ser Glu Asp Arg Ser Ile Arg
                       55
                                            60
Ser Lys Ser Ser Leu Gly Ala Gly Phe Thr Gly Glu Ser Thr Phe Thr
                                        75
                    70
Ser Lys Val Ser Ile Gln Phe Asn Leu Thr Ser Gly Met Thr Gly Leu
                                    90
Arg Ala Ser Gly Asn Pro Ser Ile Ala Gly Phe Ser Gly Ile Ser Leu
                                105
            100
Leu Ser Gly Leu Val Ala Glu Phe Cys Leu Glu Arg Pro Gly Ser Leu
                                                125
                            120
Gly Leu Cys Ala Ile Tyr Ala Ala Trp Val Gly Gly Phe Ser Met Ser
                       135
His Arg Ser Met His Asp Phe Thr Arg
145
                    150
<210> 201
<211> 527
<212> DNA
<213> Homo sapiens
gatgtggcta ttatccctgt ttcccaggtg agaaacaggg tcagtgatag agctgggatg
tgtgcctgca ggctcaccag ccagtcccct cctcaccaag gatgatgttc tccgtggtga
120
getggteett ggteteetgg aactegtgge geacetggge cagetgegee tegaaggeat
180
cottetecat etetttgget agetgeaagt tetggagetg etegttgagg tetgtgatet
catccacctg ctggttgagc gtgcgcttga ggaaggccac aatctccttc ttgttattgg
ccagetgete aaacteetgg eggaacatet teteetgeae ageeagetea teecaettee
getggtaceg ggetageegg teetecaggt eteggatetg gatgtggtag aacteettea
totoottggc cagaggcggc tocacggcca ccaccggctc cttottgccc cotttettet
tgacttcaag ctccttgcct gccttgctca cactctttt gggaggc
527
<210> 202
```

```
<211> 70
<212> PRT
<213> Homo sapiens
<400> 202
Gly Arg Pro Gln Ser Pro Ser Cys Tyr Trp Pro Ala Ala Gln Thr Pro
Gly Gly Thr Ser Ser Pro Ala Gln Pro Ala His Pro Thr Ser Ala Gly
Thr Gly Leu Ala Gly Pro Pro Gly Leu Gly Ser Gly Cys Gly Arg Thr
Pro Ser Ser Pro Trp Pro Glu Ala Ala Pro Arg Pro Pro Pro Ala Pro
                        55
Ser Cys Pro Leu Ser Ser
<210> 203
<211> 304
<212> DNA
<213> Homo sapiens
<400> 203
ngtgcaccgg tggtcatgga caacgccgcc tacgtggtct acacctcggg atccaccggc
cgacccaagg gagttgtcgt cacccacacc ggactcgaca gcttcgcact cgaccagcag
egtegattee aegeagatea ceaetetega accetgeact tegecaecee cagettegae
ggagccgtct tcgagtacct gcaggcattc ggtgtcggag ccaccatggt gatcgtcccg
accgacatet acggcggcgc cgaactggca agteteatec geegegaaca cgteacteac
gcgt
304
<210> 204
<211> 101
<212> PRT
<213> Homo sapiens
<400> 204
Xaa Ala Pro Val Val Met Asp Asn Ala Ala Tyr Val Val Tyr Thr Ser
Gly Ser Thr Gly Arg Pro Lys Gly Val Val Val Thr His Thr Gly Leu
                                25
Asp Ser Phe Ala Leu Asp Gln Gln Arg Arg Phe His Ala Asp His His
                            40
Ser Arg Thr Leu His Phe Ala Thr Pro Ser Phe Asp Gly Ala Val Phe
                                            60
                        55
Glu Tyr Leu Gln Ala Phe Gly Val Gly Ala Thr Met Val Ile Val Pro
                                        75
Thr Asp Ile Tyr Gly Gly Ala Glu Leu Ala Ser Leu Ile Arg Arg Glu
                                    90
His Val Thr His Ala
```

100 <210> 205 <211> 356 <212> DNA <213> Homo sapiens <400> 205 nngaattcag caatgataac tggctcaatt gaaggtaaga caacaattga gggaattaat gcacaattaa atacagtgtt aactttattt tcaccacaat caaaagataa agatttaatc atqccaqatc aacaaqaaga aatagatatt ctgattgcaa ccgactgtat ttcagaagga cagaacttac aagattgtga ttacttaata aactatgaca ttcattggaa tccagttcgt atcattcaaa gatttggacg gattgatcga attggttcga agaataaatg tgtacaatta qttaactttt ggccagatat tacattagat gaatatattg atctaaaggg acgcgt <210> 206 <211> 118 <212> PRT <213> Homo sapiens <400> 206 Xaa Asn Ser Ala Met Ile Thr Gly Ser Ile Glu Gly Lys Thr Thr Ile 10 Glu Gly Ile Asn Ala Gln Leu Asn Thr Val Leu Thr Leu Phe Ser Pro 25 Gln Ser Lys Asp Lys Asp Leu Ile Met Pro Asp Gln Gln Glu Glu Ile 45 40 Asp Ile Leu Ile Ala Thr Asp Cys Ile Ser Glu Gly Gln Asn Leu Gln Asp Cys Asp Tyr Leu Ile Asn Tyr Asp Ile His Trp Asn Pro Val Arg 75 70 Ile Ile Gln Arg Phe Gly Arg Ile Asp Arg Ile Gly Ser Lys Asn Lys 90 Cys Val Gln Leu Val Asn Phe Trp Pro Asp Ile Thr Leu Asp Glu Tyr 105 100 Ile Asp Leu Lys Gly Arg 115 <210> 207 <211> 324 <212> DNA <213> Homo sapiens <400> 207 acgcgtgcac tgtgtgtatg catggtaacg tacacgtgtg cactgtgtgt ggtgtgcatg

catggtgtgt geacgtgtng cactgtgtgt ggatgcatgg taatgtgcac gtgtgcactg

120

```
tqtqtqqtqt qtatqcatqq tgtgtgcacg tgtgcactgt gtgtgtgtgt atgcatgtgt
gtgcacatgt gcactgtgtg gtgtgtatgc atggtgtgtg cacgtgtgca ctgtgtatgc
atgngtgtgt gcatgtgtgc actgtgtatg catagtgtgc acgtgtgcac tgtgtggtgt
300
gtatgcatgg taatgtgcac gtgt
<210> 208
<211> 108
<212> PRT
<213> Homo sapiens
<400> 208
Thr Arg Ala Leu Cys Val Cys Met Val Thr Tyr Thr Cys Ala Leu Cys
                                    10
Val Val Cys Met His Gly Val Cys Thr Cys Xaa Thr Val Cys Gly Cys
Met Val Met Cys Thr Cys Ala Leu Cys Val Val Cys Met His Gly Val
Cys Thr Cys Ala Leu Cys Val Cys Val Cys Met Cys Val His Met Cys
                                            60
Thr Val Trp Cys Val Cys Met Val Cys Ala Arg Val His Cys Val Cys
                    70
                                        75
Met Xaa Val Cys Met Cys Ala Leu Cys Met His Ser Val His Val Cys
                85
                                    90
Thr Val Trp Cys Val Cys Met Val Met Cys Thr Cys
           100
<210> 209
<211> 168
<212> DNA
<213> Homo sapiens
<400> 209
nnctccagag gttatgaggt tggaagcccg gtttttttca ggtgcagaaa aggctaccat
attcaaggtt ccacgactcg cacctgcctt gccaatttaa catggagtgg gatacagacc
gaatgtatac ctcatgcctg cagacagcca gaaaccccgg cacacgcg
168
<210> 210
<211> 56
<212> PRT
<213> Homo sapiens
<400> 210
Xaa Ser Arg Gly Tyr Glu Val Gly Ser Pro Val Phe Phe Arg Cys Arg
                                    10
Lys Gly Tyr His Ile Gln Gly Ser Thr Thr Arg Thr Cys Leu Ala Asn
                                25
Leu Thr Trp Ser Gly Ile Gln Thr Glu Cys Ile Pro His Ala Cys Arg
```

```
45
        35
                            40
Gln Pro Glu Thr Pro Ala His Ala
    50
<210> 211
<211> 354
<212> DNA
<213> Homo sapiens
<400> 211
tacatgggct ttgacacagt ggtggctgaa gctgcactaa gggtgtttgg aggcaatgtc
cagetggeag eteagaceet tgcacaceat ggaggaagee teecaceega eetgeagtte
teaggagagg actectede cacacegted acatecedat etgactetge agggacetet
agtgcctcga cagatgaaga catggagacg gaggctgtca acgaaatcct ggaggacatt
ccggagcacg aggaggacta cctggactcc acgctggagg atgaagaagt cattattgct
gaatacttgt cctgcgttga aagtataagt tctgccngca aagaacaact gatc
354
<210> 212
<211> 118
<212> PRT
<213> Homo sapiens
<400> 212
Tyr Met Gly Phe Asp Thr Val Val Ala Glu Ala Ala Leu Arg Val Phe
                                    10
Gly Gly Asn Val Gln Leu Ala Ala Gln Thr Leu Ala His His Gly Gly
            20
                                25
Ser Leu Pro Pro Asp Leu Gln Phe Ser Gly Glu Asp Ser Ser Pro Thr
                            40
Pro Ser Thr Ser Pro Ser Asp Ser Ala Gly Thr Ser Ser Ala Ser Thr
                        55
                                            60
Asp Glu Asp Met Glu Thr Glu Ala Val Asn Glu Ile Leu Glu Asp Ile
                                        75
Pro Glu His Glu Glu Asp Tyr Leu Asp Ser Thr Leu Glu Asp Glu Glu
                                    90
Val Ile Ile Ala Glu Tyr Leu Ser Cys Val Glu Ser Ile Ser Ser Ala
                                105
Xaa Lys Glu Gln Leu Ile
        115
<210> 213
<211> 669
<212> DNA
<213> Homo sapiens
<400> 213
attgcccaat ctcagagtgt ccaggaaagc ctggagagcc tgttgcagtc tattggggaa
60
```

```
gttgaacaaa acctggaagg gaaacaggtg tcatcactct catcaggagt catccaggaa
120
qccttaqcca caaatatgaa attgaagcag gacattgctc ggcaaaaagag cagcttggag
180
gccacccgtg agatggtgac ccgattcatg gagacagcag acagtactac agcagcagtg
ctgcagggca aactggcaga ggtgagccag cggttcgaac agctctgtct acagcagcaa
gaaaaggaga geteectaaa gaagetteta eeccaggeag agatgtttga acaeetetet
ggtaagetge ageagtteat ggaaaaeaaa agteggatge tggeetetgg aaateageea
gatcaagata ttacacattt cttccaacag atccaggagc tcaatttgga aatggaagac
caacaggaga acctagatac tettgagcac etggtcaetg aactgagete ttgtggettt
gegetggaet tgtgccagca teaggaeagg gtacagaate taagaaaaga etteacagag
600
ctacagaaga cagttaaaga gagagagaaa gatgcatcat cttgccagga acagttggat
660
gaattccgg
669
<210> 214
<211> 223
<212> PRT
<213> Homo sapiens
<400> 214
Ile Ala Gln Ser Gln Ser Val Gln Glu Ser Leu Glu Ser Leu Leu Gln
                                    10
Ser Ile Gly Glu Val Glu Gln Asn Leu Glu Gly Lys Gln Val Ser Ser
            20
                                25
Leu Ser Ser Gly Val Ile Gln Glu Ala Leu Ala Thr Asn Met Lys Leu
Lys Gln Asp Ile Ala Arg Gln Lys Ser Ser Leu Glu Ala Thr Arg Glu
                                            60
Met Val Thr Arg Phe Met Glu Thr Ala Asp Ser Thr Thr Ala Ala Val
                                        75
65
Leu Gln Gly Lys Leu Ala Glu Val Ser Gln Arg Phe Glu Gln Leu Cys
                85
Leu Gln Gln Gln Glu Lys Glu Ser Ser Leu Lys Lys Leu Leu Pro Gln
                                105
Ala Glu Met Phe Glu His Leu Ser Gly Lys Leu Gln Gln Phe Met Glu
                            120
        115
Asn Lys Ser Arg Met Leu Ala Ser Gly Asn Gln Pro Asp Gln Asp Ile
                        135
                                            140
Thr His Phe Phe Gln Gln Ile Gln Glu Leu Asn Leu Glu Met Glu Asp
                                        155
                    150
Gln Gln Glu Asn Leu Asp Thr Leu Glu His Leu Val Thr Glu Leu Ser
                                    170
Ser Cys Gly Phe Ala Leu Asp Leu Cys Gln His Gln Asp Arg Val Gln
                                185
Asn Leu Arg Lys Asp Phe Thr Glu Leu Gln Lys Thr Val Lys Glu Arg
```

```
195
                            200
Glu Lys Asp Ala Ser Ser Cys Gln Glu Gln Leu Asp Glu Phe Arg
    210
                        215
<210> 215
<211> 814
<212> DNA
<213> Homo sapiens
<400> 215
aaatttegta eeegeteegg cacagtaega geeettgaeg atgtgageet ggetattaag
agaggttcca totcagccgt tatcgggcac tccggagccg gcaaatccac cctggttcgc
ctcatcaacq gattaqagac tcccacgcgt ggccgcgtct tggtagacgg caccgacgtc
tegeagetet eggacaaage gatgegeeeg etaegegeag acategggat gatetteeaa
cagttcaacc tattcggctc aaggaccatc tacgacaacg ttgcctatcc actcaagctg
geteattgga agaaageaga egagaagaag egegteaeeg aattgetgag ettegteggg
360
ttgacgagca aagcctggga ccatccagac cagctctcgg gcggacagaa acagcgggtt
ggtattgccc gagcgctagc aactaaacca tcgattttgt tggctgacga gtccacctcg
gegetggate cagaaacgac agetgatgte etatecetge teaagegggt caatgeggaa
ctaqqqqtqa cqgtcqtcqt catcacccac gagatggagg tcgtccgctc gattgcccag
caggicticg tactagoago tggccatoto giogagictig gaagogooog coaggictic
gctcatccac agtcagagac cacccagcgt ttcctggcga cgattatcgg ccagcacccg
agtggggagg aacaggcacg gttgcagtcg gaaaacccag atgcacgact cgtcgacgtc
agttcggtgg ccagtcactc gttcggtgac gcgt
814
<210> 216
<211> 271
<212> PRT
<213> Homo sapiens
<400> 216
Lys Phe Arg Thr Arg Ser Gly Thr Val Arg Ala Leu Asp Asp Val Ser
Leu Ala Ile Lys Arg Gly Ser Ile Ser Ala Val Ile Gly His Ser Gly
            20
                                25
Ala Gly Lys Ser Thr Leu Val Arg Leu Ile Asn Gly Leu Glu Thr Pro
                            40
                                                45
Thr Arg Gly Arg Val Leu Val Asp Gly Thr Asp Val Ser Gln Leu Ser
                        55
Asp Lys Ala Met Arg Pro Leu Arg Ala Asp Ile Gly Met Ile Phe Gln
```

```
70
Gln Phe Asn Leu Phe Gly Ser Arg Thr Ile Tyr Asp Asn Val Ala Tyr
                                    90
Pro Leu Lys Leu Ala His Trp Lys Lys Ala Asp Glu Lys Lys Arg Val
                                105
Thr Glu Leu Leu Ser Phe Val Gly Leu Thr Ser Lys Ala Trp Asp His
                            120
Pro Asp Gln Leu Ser Gly Gly Gln Lys Gln Arg Val Gly Ile Ala Arg
                       135
Ala Leu Ala Thr Lys Pro Ser Ile Leu Leu Ala Asp Glu Ser Thr Ser
                                        155
                    150
Ala Leu Asp Pro Glu Thr Thr Ala Asp Val Leu Ser Leu Leu Lys Arg
                165
                                    170
Val Asn Ala Glu Leu Gly Val Thr Val Val Val Ile Thr His Glu Met
                                185
            180
Glu Val Val Arg Ser Ile Ala Gln Gln Val Ser Val Leu Ala Ala Gly
                            200
His Leu Val Glu Ser Gly Ser Ala Arg Gln Val Phe Ala His Pro Gln
                                            220
                        215
Ser Glu Thr Thr Gln Arg Phe Leu Ala Thr Ile Ile Gly Gln His Pro
                    230
                                        235
Ser Gly Glu Glu Gln Ala Arg Leu Gln Ser Glu Asn Pro Asp Ala Arg
                                   250
               245
Leu Val Asp Val Ser Ser Val Ala Ser His Ser Phe Gly Asp Ala
                                265
<210> 217
<211> 500
<212> DNA
<213> Homo sapiens
<400> 217
nnacgcgtcg cgatgaaaga ggcgctgaaa ggtgccatcc agattccaac agtgactttt
agetetgaga agtecaatac tacagecetg getgagtteg gaaaatacat teataaagte
tttcctacag tggtcagcac cagctttatc cagcatgaag tcgtggaaga gtatagccac
ctgttcacta tccaaggctc ggaccccage ttgcagccct acctgctgat ggctcacttt
gatgtggtgc ctgcccctga agaaggctgg gaggtgcccc cattctctgg gttggagcgt
300
gatggcgtca totatggttg gggcacactg gacgacaaga actotgtgat ggcattactg
caggeettgg ageteetget gateaggaag tacateecee gaagatettt etteatttet
420
ctgggccatg atgaggagtc atcagggaca ggggctcaga ggatctcagc cctgctacag
tcaaggggcg tccagctagc
500
<210> 218
<211> 166
<212> PRT
```

<213> Homo sapiens <400> 218 Xaa Arg Val Ala Met Lys Glu Ala Leu Lys Gly Ala Ile Gln Ile Pro Thr Val Thr Phe Ser Ser Glu Lys Ser Asn Thr Thr Ala Leu Ala Glu 25 Phe Gly Lys Tyr Ile His Lys Val Phe Pro Thr Val Val Ser Thr Ser 40 Phe Ile Gln His Glu Val Val Glu Glu Tyr Ser His Leu Phe Thr Ile 55 Gln Gly Ser Asp Pro Ser Leu Gln Pro Tyr Leu Leu Met Ala His Phe 70 75 Asp Val Val Pro Ala Pro Glu Glu Gly Trp Glu Val Pro Pro Phe Ser 90 85 Gly Leu Glu Arg Asp Gly Val Ile Tyr Gly Trp Gly Thr Leu Asp Asp 105 Lys Asn Ser Val Met Ala Leu Leu Gln Ala Leu Glu Leu Leu Ile 120 Arg Lys Tyr Ile Pro Arg Arg Ser Phe Phe Ile Ser Leu Gly His Asp 135 140 Glu Glu Ser Ser Gly Thr Gly Ala Gln Arg Ile Ser Ala Leu Leu Gln 155 Ser Arg Gly Val Gln Leu 165 <210> 219 <211> 361 <212> DNA <213> Homo sapiens <400> 219 acgcgttgaa acgggtatat tggggatgac gccgctgtgc aatatgcgca aggccataca caaggteege aegeteecat gteectegtt ttegacagtt ettttgegee geattatgge gaageegteg agattgegee tgatateaag egeateaegg teaacaacee eageeeette 180 actttttcg gcaccaacag ttatctgatc ggccgcgata cgctggcatt gatcgatccc ggtccgcttg acgaggccca tcacgcggcg ctgctgcgtg ccattgccgg ccggccggtc agccatatet ttgteageea cacacacegg gaccactege cagtegegae ggttttgaaa 360 g 361 <210> 220 <211> 102 <212> PRT <213> Homo sapiens Met Ala Asp Arg Pro Ala Gly Asn Gly Thr Gln Gln Arg Arg Val Met

```
Gly Leu Val Lys Arg Thr Gly Ile Asp Gln Cys Gln Arg Ile Ala Ala
Asp Gln Ile Thr Val Gly Ala Glu Lys Ser Glu Gly Ala Gly Val Val
                            40
Asp Arg Asp Ala Leu Asp Ile Arg Arg Asn Leu Asp Gly Phe Ala Ile
                       55
Met Arg Arg Lys Arg Thr Val Glu Asn Glu Gly His Gly Ser Val Arg
                                        75
                    70
Thr Leu Cys Met Ala Leu Arg Ile Leu His Ser Gly Val Ile Pro Asn
                                   90
Ile Pro Val Ser Thr Arg
           100
<210> 221
<211> 401
<212> DNA
<213> Homo sapiens
<400> 221
agatetetgt gtegtegget geaaagagga tgageecaga tgcatateag gggeteecte
ccacatccca cctgctcggg cagcccacgg cagccccaca ctgctgcagc acacctcgct
quaqctotqq ttoctoctca quantatocc tgccaccotg ctaagcottg gccaacactg
caccetgtee caatgegget ceagtgacea cacceceagg geataceete etacagagea
ttcccaaaaa aggctagagt agacaccagc ctgctccgta gggggcctcc accccattct
ccaaggeete cacccaggga cgcctggtga accagcatee aggeetggee caccteeetg
ctcagagtcc atgttctgtg acaagggtgg caactgggat t
<210> 222
<211> 124
<212> PRT
<213> Homo sapiens
<400> 222
Met Asp Ser Glu Gln Gly Gly Pro Gly Leu Asp Ala Gly Ser Pro
                                   10
Gly Val Pro Gly Trp Arg Pro Trp Arg Met Gly Trp Arg Pro Pro Thr
                                25
Glu Gln Ala Gly Val Tyr Ser Ser Leu Phe Trp Glu Cys Ser Val Gly
                            40
Gly Tyr Ala Leu Gly Val Trp Ser Leu Glu Pro His Trp Asp Arg Val
                        55
                                            60
Gln Cys Trp Pro Arg Leu Ser Arg Val Ala Gly Ile Phe Leu Arg Arg
                                        75
Asn Gln Ser Cys Ser Glu Val Cys Cys Ser Ser Val Gly Leu Pro Trp
Ala Ala Arg Ala Gly Gly Met Trp Glu Gly Ala Pro Asp Met His Leu
```

```
110
            100
                                105
Gly Ser Ser Ser Leu Gln Pro Thr Thr Gln Arg Ser
                            120
<210> 223
<211> 331
<212> DNA
<213> Homo sapiens
<400> 223
tcatgaaatc tgtgggcagt gacccaggag ggtatgggca ggcccaacca ggttggtgtg
cccttgaagc cccacagacc tgccagggca gcagggcagt tgggagccgg agaacctgag
aaccaagcca ggctgcatgc aggaggctgg cacgtgaacg ctgcaggtgt tgccggcagc
egtggtgcct ggcagatagt gttcgacccc cnaggacctt cttgctgggc agcccagtcc
240
aaaagetgtt cccgcttaag ccaccccac cgccttggcc acacctggca catgggtgaa
gcaagggcat ttcccggggc ttcctgttcc c
331
<210> 224
<211> 103
<212> PRT
<213> Homo sapiens
<400> 224
Met Pro Leu Leu His Pro Cys Ala Arg Cys Gly Gln Gly Gly Gly Gly
                                    10
Gly Leu Ser Gly Asn Ser Phe Trp Thr Gly Leu Pro Ser Lys Lys Val
            20
                                25
                                                     3.0
Leu Gly Gly Arg Thr Leu Ser Ala Arg His His Gly Cys Arg Gln His
Leu Gln Arg Ser Arg Ala Ser Leu Leu His Ala Ala Trp Leu Gly Ser
                        55
Gln Val Leu Arg Leu Pro Thr Ala Leu Leu Pro Trp Gln Val Cys Gly
                    70
                                        75
Ala Ser Arg Ala His Gln Pro Gly Trp Ala Cys Pro Tyr Pro Pro Gly
                85
                                    90
Ser Leu Pro Thr Asp Phe Met
            100
<210> 225
<211> 339
<212> DNA
<213> Homo sapiens
<400> 225
tgatcacggg cgtgagccac cagcccagca tcccttgcct ttcattcgca cctccacctc
cagaatgacc ctcattccct cctgcacaga cggtgacagc agtaactcct acaaacacca
120
```

```
ccagactgat cttcaagagc agaggaactc ccaatcacga ttccaccccc gccgggctct
180
caaatcctcc agggctgcct gctatggggg agggaggcac actttgcttg gctctcaagg
240
ceteagecag eegggteeaa accaacteee ageetggeet caccateeca eegeeaaace
tttqctcaca ctggcccctc ttcctggaac atgggcctn
339
<210> 226
<211> 91
<212> PRT
<213> Homo sapiens
<400> 226
Met Thr Leu Ile Pro Ser Cys Thr Asp Gly Asp Ser Ser Asn Ser Tyr
                                    10
Lys His His Gln Thr Asp Leu Gln Glu Gln Arg Asn Ser Gln Ser Arg
Phe His Pro Arg Arg Ala Leu Lys Ser Ser Arg Ala Ala Cys Tyr Gly
                            40
Gly Gly Arg His Thr Leu Leu Gly Ser Gln Gly Leu Ser Gln Pro Gly
                        55
                                            60
Pro Asn Gln Leu Pro Ala Trp Pro His His Pro Thr Ala Lys Pro Leu
                    70
Leu Thr Leu Ala Pro Leu Pro Gly Thr Trp Ala
                85
<210> 227
<211> 353
<212> DNA
<213> Homo sapiens
<400> 227
qtcqaccct tcgattgtgg cgaactccat ggctgctgcg ggcctgcgta ggctctcgag
tagetegacg tegggttege gagggetege agegtggeea tgetgettet tggatggtte
120
gggcaactcc tcgggggatt cgagcagttc ttggcgcacc tgctctggcg tcatcccgga
ggccaggccg acaagtgctg cctcctgcca cccgctgagc gacgctgcca tgttgagtac
ggegtettea etggteaggg egagegeggt ategaceagg ttggegteea ggeegagaga
cagcatgtct gctcagtcgc ggtgatgact ggagtggcgg tctcctgcac ggg
353
<210> 228
<211> 102
<212> PRT
<213> Homo sapiens
Met Leu Ser Leu Gly Leu Asp Ala Asn Leu Val Asp Thr Ala Leu Ala
```